VOLUME 37 NOVEMBER 1959 NUMBER 11

Canadian Journal of Physics

Editor: H. E. DUCKWORTH

Associate Editors:

L. G. ELLIOTT, *Atomic Energy of Canada, Ltd., Chalk River*
J. S. FOSTER, *McGill University*
G. HERZBERG, *National Research Council of Canada*
L. LEPRINCE-RINGUET, *Ecole Polytechnique, Paris*
B. W. SARGENT, *Queen's University*
G. M. VOLKOFF, *University of British Columbia*
W. H. WATSON, *University of Toronto*
G. A. WOONTON, *McGill University*

Published by THE NATIONAL RESEARCH COUNCIL

OTTAWA CANADA

CANADIAN JOURNAL OF PHYSICS

Under the authority of the Chairman of the Committee of the Privy Council on Scientific and Industrial Research, the National Research Council issues THE CANADIAN JOURNAL OF PHYSICS and five other journals devoted to the publication, in English or French, of the results of original scientific research. Matters of general policy concerning these journals are the responsibility of a joint Editorial Board consisting of: members representing the National Research Council of Canada; the Editors of the Journals; and members representing the Royal Society of Canada and four other scientific societies.

EDITORIAL BOARD

Representatives of the National Research Council

I. McT. Cowan, *University of British Columbia*
A. Gauthier, *University of Montreal*
H. G. Thode (Chairman), *McMaster University*
D. L. Thomson, *McGill University*

Editors of the Journals

D. L. Bailey, *University of Toronto*
T. W. M. Cameron, *Macdonald College*
H. E. Duckworth, *McMaster University*
Léo Marion, *National Research Council*
J. F. Morgan, *Department of National Health and Welfare, Ottawa*
R. G. E. Murray, *University of Western Ontario*
J. A. F. Stevenson, *University of Western Ontario*

Representatives of Societies

D. L. Bailey, *University of Toronto*
Royal Society of Canada
T. W. M. Cameron, *Macdonald College*
Royal Society of Canada
H. E. Duckworth, *McMaster University*
Royal Society of Canada
Canadian Association of Physicists
P. R. Gendron, *University of Ottawa*
Chemical Institute of Canada
J. F. Morgan, *Department of National Health and Welfare, Ottawa*
Canadian Biochemical Society
R. G. E. Murray, *University of Western Ontario*
Canadian Society of Microbiologists
J. A. F. Stevenson, *University of Western Ontario*
Canadian Physiological Society
T. Thorvaldson, *University of Saskatchewan*
Royal Society of Canada

Ex officio

Léo Marion (Editor-in-Chief), *National Research Council*
J. B. Marshall (Administration and Awards), *National Research Council*

Manuscripts for publication should be submitted to Dr. H. E. Duckworth, Editor, Canadian Journal of Physics, Hamilton College, McMaster University, Hamilton, Ontario.
For instructions on preparation of copy, see **NOTES TO CONTRIBUTORS** (back cover).

Proof, correspondence concerning proof, and orders for reprints should be sent to the Manager, Editorial Office (Research Journals), Division of Administration and Awards, National Research Council, Ottawa 2, Canada.

Subscriptions, renewals, requests for single or back numbers, and all remittances should be sent to Division of Administration and Awards, National Research Council, Ottawa 2, Canada. Remittances should be made payable to the Receiver General of Canada, credit National Research Council.

The journals published, frequency of publication, and prices are:

Canadian Journal of Biochemistry and Physiology	Monthly	$9.00 a year
Canadian Journal of Botany	Bimonthly	$6.00 a year
Canadian Journal of Chemistry	Monthly	$12.00 a year
Canadian Journal of Microbiology	Bimonthly	$6.00 a year
Canadian Journal of Physics	Monthly	$9.00 a year
Canadian Journal of Zoology	Bimonthly	$5.00 a year

The price of regular single numbers of all journals is $2.00.

Canadian Journal of Physics

Issued by THE NATIONAL RESEARCH COUNCIL OF CANADA

VOLUME 37 NOVEMBER 1959 NUMBER 11

THE SUN AS A SOURCE OF COSMIC RAYS OF INTERMEDIATE ENERGIES[1]

J. KATZMAN

ABSTRACT

The cosmic ray intensity as measured with an extremely narrow-angle telescope, 1.2×10^{-3} steradians, and with 96 inches of lead as absorber for the period 1 January 1955 to 31 December 1958 shows an increase of 20%. This increase is attributed to particles coming from the sun. It is shown that the change in hour of maximum of the first and second harmonics can be explained by a change in the relative importance of the impact zones. This phenomenon is attributed to a change in the number and polarity of sunspots.

INTRODUCTION

The long-term change in the intensity of cosmic rays has been shown by Forbush (1954), Fenton *et al.* (1958), and others to relate inversely with solar activity. The activity has a pronounced influence on the intensity of particles of low energies as monitored by neutron monitors, ionization chambers, and Geiger counter telescopes. The variation of cosmic rays of high energies is studied by means of extensive air showers but such experiments have not operated continuously for a complete solar cycle. To avoid confusion we will call "intermediate energies" those energies of particles between that monitored by Geiger counter telescopes with the conventional 10 cm of lead as absorber and those energies of particles studied by air showers. The experiment by Rau (1939), who lowered ionization chambers into Lake Constance to a depth of 40 meters during the solar maximum of 1937–38, will fall into this category. The very interesting fact he reported, though unexplained, was that the second harmonic of the cosmic ray intensity showed a positive correlation with the second harmonic of atmospheric pressure.

The activity of the sun is known to pass from minimum to maximum in a period of approximately 11 years and is measured in terms of the number of sunspots, the magnetic polarity of which has a period of approximately 22 years (The Sun 1953). Thambyahpillai and Elliot (1953) have suggested that the change in hour of maximum of the daily cosmic ray intensity may be related to the 22-year sunspot cycle. Since the sun is an active body and is

[1]Manuscript received August 6, 1959.
Contribution from the Division of Pure Physics, National Research Council, Ottawa. Canada.
Issued as N.R.C. No. 5370.

continuously emitting energy in one form or another, it is reasonable to consider that the sun under certain conditions does emit particles. When the activity is at a low ebb, as measured in a solar cycle, the particles will be of low energy and will either not leave the influence of the sun or those that do reach the outer atmosphere of the earth will not have sufficient energy so that their progeny will reach the surface of the earth. In this study we consider the solar minimum activity as that period when particles from the sun do not have sufficient energy to influence monitors on the earth and the intensity of cosmic rays during this period is that of particles that come from outside the solar system.

The magnetic field that surrounds the earth is more that of a dipole at some distance from the earth than at comparatively small distances from the earth. The field that a cosmic ray particle encounters is not that of the earth's field only (Rose *et al.* 1956). It has been shown by Dwight (1950), Firor (1954), Lüst (1957, 1958) that if the magnetic field is considered to be that of a dipole then certain hours of the day (impact zones) will show a greater intensity of cosmic rays than the other hours if particles are arriving from the vicinity of the sun. They found 0300, 0900, and 2000 to be the important hours. A radical change in the magnetic field that the cosmic ray particles see as they approach the earth will change the relative importance of the impact zones.

To explain the sudden decreases in intensity of cosmic rays known as Forbush decreases, which last for a number of days, and the long-term 11-year cycle a "barrier" is considered to be created by the sun's activity that bars, either by electrical or magnetic means, particles of low energy from reaching the earth (Simpson 1957, 1958; Morrison 1956; Alfven 1954; and Parker 1956). This barrier must be maintained and intensified with solar activity to explain the 11-year cycle in cosmic ray intensity. To explain the Forbush decrease the barrier must be increased suddenly for a short period. No field of force can exist unless it is associated with matter in one form or another. If we accept the hypothesis that the sun emits particles at all times and that a barrier is created by this activity then the following can be expected:

(1) a net decrease in low-energy particles as monitored by neutron monitors, if the barrier stops more low-energy particles from reaching the earth than the sun emits;

(2) a net decrease in particles as monitored by Geiger counter telescopes or ionization chambers with 10 cm of lead as absorber as the barrier intensity increases;

(3) an increase in intensity of short duration before a Forbush decrease, the intensity of the barrier being increased after the particles pass the earth;

(4) a possible increase in intensity as monitored by telescopes not sensitive to particles of energies below the intermediate range.

EXPERIMENTAL

A narrow-angle telescope (1.2×10^{-3} steradians) with 96 inches of lead as absorber has been operating continuously since December 1954 (Katzman

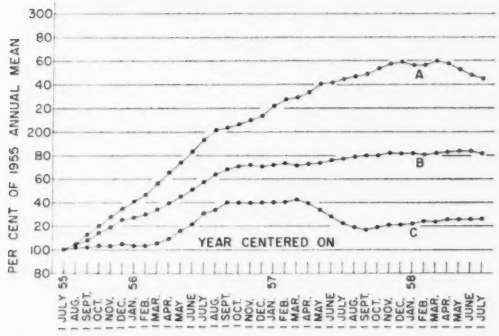


FIG. 1. Twelve-month running averages in percentage of the 1955 annual mean of (A) the solar flux (2800 Mc/sec), (B) the critical frequency of the F_2 layer, (C) the cosmic ray intensity as measured by the narrow-angle telescope.

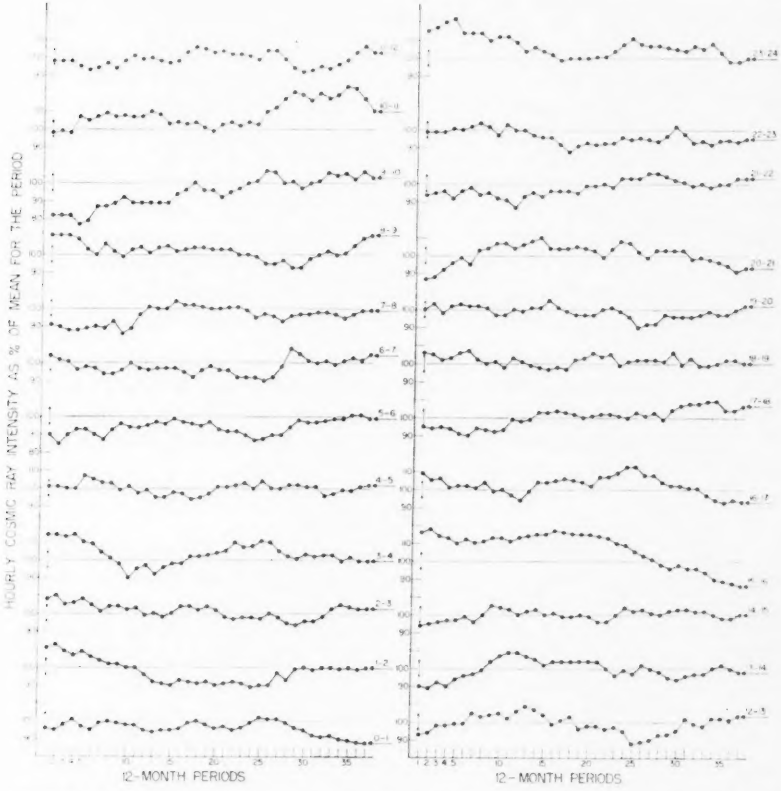


FIG. 2. Twelve-month running averages of the annual hourly intensity in percentage of the 12-month period.

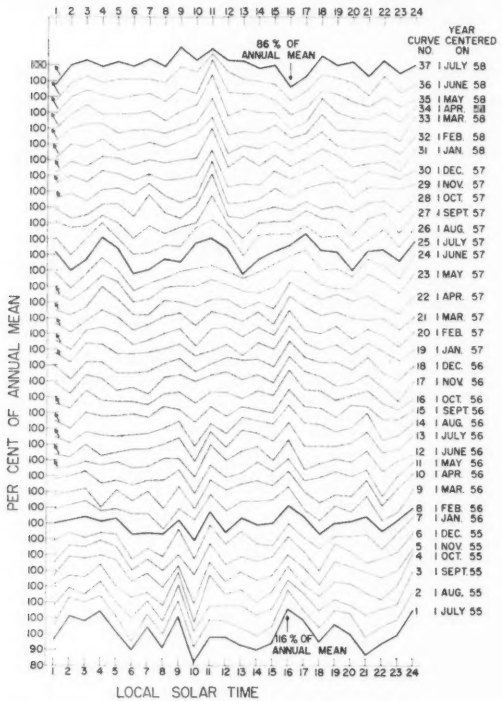


FIG. 3. Twelve-month running averages of the annual daily intensity in percentage of the 12-month period.

1958). Triple coincidences are recorded on a 20-pen Esterline Angus tape recorder and on scalers. The data, for the period 1 January 1956 to 31 December 1958, were analyzed and are presented in the form of curves (Figs. 1–3). The data were not corrected for either temperature or pressure and are shown as 12-month running averages.

RESULTS

The 2800 Mc/sec radio wave, emitted from the sun, is used as a measure of solar activity (Covington 1957). The curves in Fig. 1 are 12-month running averages in percentage of the 1955 average value. Curve A is a plot of the radiation intensity of a frequency of 2800 Mc/sec,[2] curve B of the critical frequency of the F_2 layer,[3] and curve C of the intensity of cosmic rays as monitored by the narrow-angle telescope. A 12-month period of data is referred to by the first day of the seventh month of that period, thus 1 July

[2]Solar flux at 2800 Mc/sec in watts/m^2 cycles sec (bandwidth$\times 10^{-22}$). Data supplied by Mr. A. E. Covington, National Research Council, Ottawa, Ontario, Canada.

[3]Data supplied by Defence Research Telecommunications Establishment, Defence Research Board, Ottawa, Ontario, Canada.

1955 refers to the 12-month period 1 January to 31 December 1955. There is a marked change in slope of curve A at 1 August 1956, curve B at 1 October 1956, and of curve C at 1 September 1956. It is clear that at about 1 September 1956 a radical change took place in the conditions existing in the sun–earth system. The rate of change of the solar flux then decreased. The critical frequency of the F_2 layer practically stopped increasing and the cosmic ray intensity remained constant until 1 March 1957, then dropped sharply to 1 August 1957, and later rose slightly.

The cosmic ray intensity data are plotted on an hourly basis in Figs. 2 and 3. In Fig. 2 each hour of the day is plotted as a separate curve, the hour for 12-month running averages taken as percentage of the mean for that 12-month period. The 9–10 hour is only 80% of the annual mean for 1955 and rises to 105% for 1958, whereas the 15–16 hour is 116% of the 1955 mean and drops to 86% of the 1958 mean. The daily intensity is plotted as percentage of the annual mean in Fig. 3; again each curve represents a year of data centered about the day indicated. The intensity about the 0300, at the 0900, and at the 1600 and 2400 hours is above the annual mean. The pattern of curve No. 1 is maintained to curve No. 8 when there is a marked change and the pattern changes again at curve No. 25.

It can readily be seen from the curves in Fig. 3 that the hours of maximum for the first and second harmonics obtained from curve No. 1 would not be the same as that of curve No. 37. Since the 15–16 hour intensity is consistently high up to the 20th period (referring now to Fig. 2) and then reduces uniformly to a low value, the month when the value reached 100% of the annual mean was chosen as the dividing month. The data were thus divided into the two periods: 1 January 1955 to 31 August 1957 and 1 September 1957 to 31 December 1958. A harmonic analysis for the first and second components of the diurnal variation for these two periods is shown in Fig. 4. The hour

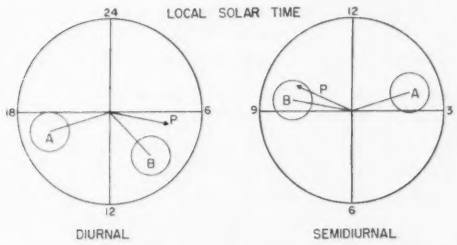


FIG. 4. Harmonic dials of the diurnal and semidiurnal components of the atmospheric pressure, P, at Ottawa, Canada, and of the cosmic ray intensity for the two periods, A and B, 1 January 1955 to 31 August 1957 and 1 September 1957 to 31 December 1958.

of maximum of the first harmonic is 16 hours 48 minutes for the first period "A" and 9 hours 8 minutes for the second period "B". The second harmonic also shows a large change, the times of maximum being 2 hours 26 minutes and 9 hours 22 minutes for A and B. It was mentioned above that these data

were not corrected for barometric pressure. A harmonic analysis of the diurnal variation in barometric pressure does not show any appreciable change over the period. Temperature effects in the atmosphere though they have a seasonal effect could not account for this long period change. It is clear, therefore, that this change in phase is not of atmospheric origin. In fact a correlation with atmospheric pressure would show a change from negative to positive. The pressure vectors "P" are also shown in Fig. 4.

DISCUSSION

The increase in ionization of the F_2 layer cannot be attributed only to an increase in electromagnetic energy emitted by the sun, since the percentage of the increase over the earth is not uniform (Katzman 1958). Therefore, some of the increase in ionization must be due to an increase in the number of particles arriving at the earth. Since it has been shown by others that the sun's activity bars low-energy particles from arriving at the earth then this increase in particles must be coming from within the sun–earth system.

The intensity of cosmic rays of intermediate energy remained constant up to 1 February 1956 and increased by 40% to 1 September 1956, remained steady till 1 April 1957 when it dropped to a level 20% above the 1955 annual mean. This increase cannot be attributed to an increase in the cosmic ray intensity outside the barrier created by the sun. The increase must, therefore, be attributed to some mechanism within the barrier. Any mechanism which would increase the intensity of intermediate energy particles would also increase the intensity of the low-energy particles. This can also be seen in the nucleon component at Ottawa from June to December 1956 in Fig. 3 of Fenton *et al.* (1958). The sunspot number continued to increase during this period whereas the nucleon component showed an increase instead of the expected decrease. Hence, since neither the sunspot number nor the solar flux showed a decrease the increase in the cosmic ray intensity up to intermediate energies for this period must be attributed to some mechanism on the sun.

Under certain conditions it is possible for the sun's activity to be the cause of an increase in cosmic ray particles of energies above 15 Bev as was evidenced by the sudden increase of 23 February 1956 at Huancayo (Simpson 1958). To study this behavior the intensity of the nucleon and meson components at Ottawa is shown as percentage of the average intensity of the first 10 days of January 1956, for the months January, February, and March (Fig. 5). There is no reason to believe that the weather conditions for similar months of other years will be so drastically different that the intensity would be radically different because of the weather; thus similar periods for the years 1957 and 1958 are also shown for comparison. These curves are very interesting. Before the "flare increase" the meson component remained above 100% for about 20 days while the nucleon component after a small decrease slowly increased in intensity to a peaked value equal to the meson component just before the Forbush-type decrease. The decrease in the nucleon component preceded the decrease in the meson component and the reverse

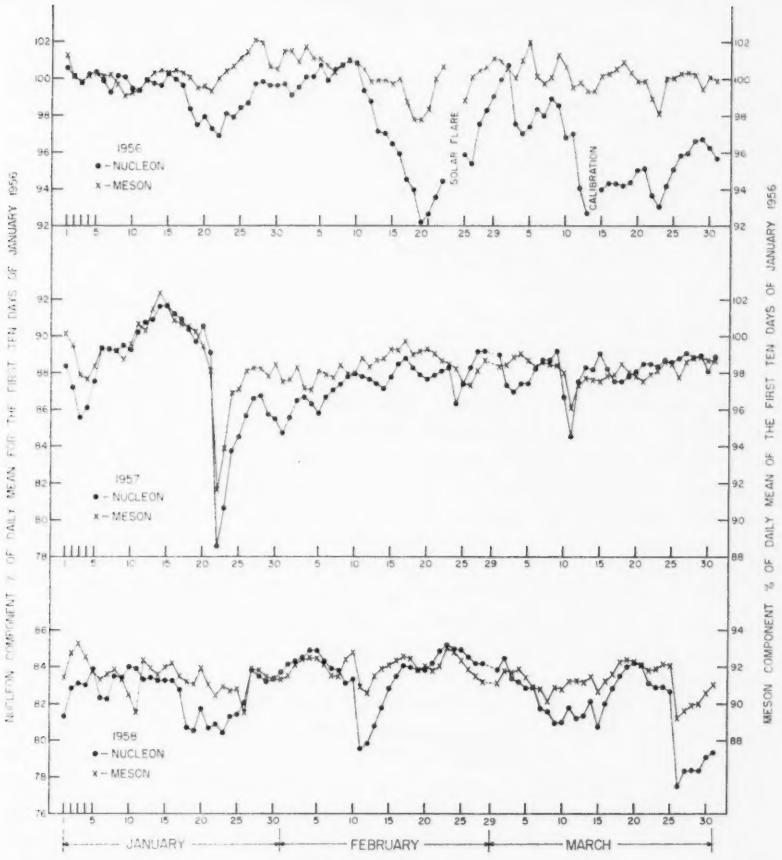


FIG. 5. The intensity at Ottawa of the nucleon and meson components, corrected for pressure, in percentage of the first 10 days of January 1956.

is obvious for the return to "normal" intensity. The following decrease in March is not as marked for the meson component but the increase before the decrease is clear in the meson component. The intense decrease of 22 January 1957 is preceded by an increase in both the nucleon and meson components centered about 14 January. If the intensity of the barrier is increasing so that particles of higher energy are being stopped from entering the sun–earth system, then with time a decrease in particles of energies between the low energy of the nucleon component and those of intermediate energies is maintained. This can be seen for the years 1957 and 1958, Figs. 5 and 6, when the intensity of the meson component dropped relatively more than the nucleon component. Since the intensity of the cosmic rays of intermediate

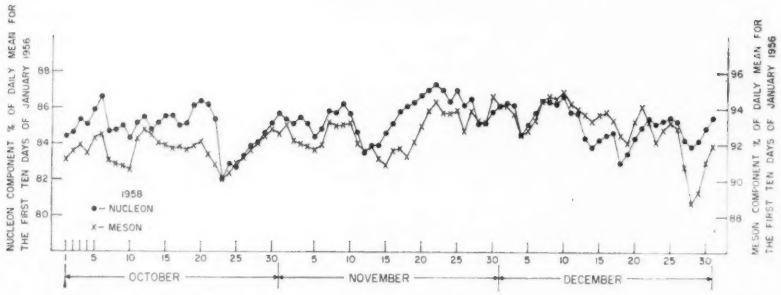

NUCLEON COMPONENT % OF DAILY MEAN FOR THE FIRST TEN DAYS OF JANUARY 1956
MESON COMPONENT % OF DAILY MEAN FOR THE FIRST TEN DAYS OF JANUARY 1956
1958
● NUCLEON
x MESON
OCTOBER
NOVEMBER
DECEMBER

FIG. 6. The intensity at Ottawa of the nucleon and meson components, corrected for pressure, in percentage of the first 10 days of January 1956 for the months October, November, and December 1958.

energies increased by about 20% while the intensity of particles of lower energies decreased it is proposed that the sun by some means as yet unexplainable causes a barrier, as suggested by others, to be created that stops cosmic ray particles of lower energies from outside the system from reaching the earth, while at the same time emitting particles of energies up to intermediate energies. Thus monitors that are sensitive to lower energies show a decrease whereas monitors that require higher energies show an increase.

The experiment by Rau (1939) at Lake Constance was performed at approximately the same geographic latitude as Ottawa. The hour of maximum for the second harmonic of atmospheric pressure would, therefore, be at about the same local time for the two stations. Since he obtained a positive correlation with pressure for cosmic rays of intermediate energies (the hour of maximum for the second harmonic was at 1000 hours) as is similarly reported in this paper it is reasonable to conclude that the cause that increased the intensity of the morning hours and decreased the intensity of the afternoon hours was the same for both the solar maximum of 1937–38 and of 1957–58. The condition on the sun that is known to be the same for these two maxima is the magnetic polarity of the sunspots. This indicates that the magnetic intensity and polarity of the sunspots move from the sun towards the earth carried by particles emitted from the sun. This influence would be greatest at the sun's maximum activity since it is then that the greatest number of spots exist on the sun. As the sun's activity decreases the number of spots decrease and the influence of the polarity of the spots decreases till eventually the earth's field becomes dominant. Rerurkar (1957) with the aid of Alfven's (1954) theory of a frozen-in magnetic field in an ionized beam emitted by the sun has shown that the hour of maximum will shift from the afternoon to morning hours when the direction of the magnetic field rotates through 180°. This is clearly shown in Figs. 2 and 3 where the predominance of the 16th hour intensity shifts with increase of solar activity to the 11th hour.

CONCLUSION

From the data presented it is concluded that the sun emits cosmic ray particles whose energy depends on the intensity of the solar activity. The magnetic field of the sunspots is carried towards the earth by the emitted particles. The change in number and polarity of the spots will influence the impact zones on the surface of the earth of cosmic ray particles and will cause an apparent change in correlation with atmospheric pressure from negative to positive. This will also cause a change in apparent anisotropy with change in solar activity.

ACKNOWLEDGMENTS

The author wishes to thank Dr. D. C. Rose for his critical review of the manuscript, Mr. O. Poulin and Mr. J. Winfield for the excellent manner in which they maintain the equipment, and the other members of the Cosmic Ray Group in the daily chore of taking readings.

REFERENCES

ALFVEN, H. 1954. Tellus, **6**, 232.
COVINGTON, A. E. 1957. J. Roy. Astron. Soc. Can. **51**, 298.
DWIGHT, K. 1950. Phys. Rev. **78**, 40.
FENTON, A. G., FENTON, K. B., and ROSE, D. C. 1958. Can. J. Phys. **36**, 824.
FIROR, J. 1954. Phys. Rev. **94**, 1017.
FORBUSH, S. E. 1954. J. Geophys. Research, **59**, 525.
KATZMAN, J. 1958. Can. J. Phys. **36**, 807.
KIEPENHEUER, K. O. 1953. Solar activity. *In* The sun, *edited by* G. P. Kuiper (The University of Chicago Press, Chicago, Illinois).
LÜST, R. and SIMPSON, J. A. 1957. Phys. Rev. **108**, 1563.
LÜST, R. 1958. Electromagnetic phenomena in cosmical physics (Cambridge University Press, London).
MORRISON, P. 1956. Phys. Rev. **101**, 1397.
PARKER, E. N. 1956. Phys. Rev. **103**, 1518.
RAU, W. 1939. Z. Phys. **114**, 265.
RERURKAR, N. W. 1957. Proc. Indiana Acad. Sci. **45**, 341.
ROSE, D. C., FENTON, K. B., KATZMAN, J., and SIMPSON, J. A. 1956. Can. J. Phys. **34**, 968.
SIMPSON, J. A. 1957. Proc. Natl. Acad. Sci. **43**, 42.
——— 1958. Electromagnetic phenomena in cosmical physics (Cambridge University Press, London).
THAMBYAHPILLAI, T. and ELLIOT, H. 1953. Nature, **171**, 918.

RESOLVING POWER OF THREE ANTENNA PATTERNS DERIVED FROM THE SAME APERTURE[1]

A. E. Covington and Gladys A. Harvey

ABSTRACT

Three antenna patterns are derived from the same linear aperture and may be described in terms of an angular spectrum of spatial frequencies ranging from zero to a common cutoff frequency. The band passes according to the shape of the spectrum are rectangular, triangular, and cosinusoidal for the three patterns, and give resolving powers respectively of 1.33, 1.00, and 1.05, in terms of the cutoff period. The rectangular band pass gives rise to the optimum antenna pattern and allows the Fourier components of a source from zero to cutoff frequency to be received with equal intensity and zero phase shift. Scanning curves of two equally intense point sources and a uniformly bright line are investigated.

INTRODUCTION

In radio astronomy, the abundance and closeness of radio-emissive objects in the sky has necessitated the construction of various types of antennas with large dimensions, in an effort to isolate and to study the size of individual objects. Although the apertures of these antennas, now in use or under construction, are measured in terms of several hundreds of meters, 1 minute of arc is about the narrowest beam width which can be realized. There are objects in the sky which are much wider than this, but other sources such as certain sunspots and radio stars may have widths of the same order. Thus, in the latter case, it is necessary to attempt to obtain a true picture of the structure of the object under examination, or at least some estimate of its structure, in the face of a serious limitation in antenna resolving power. This problem has two aspects: first, that of utilizing a given physical area to obtain an antenna pattern with the sharpest main lobe; and second, that of interpreting the departures of the observed scanning curves from the known antenna pattern in terms of object structure. Both of these aspects will be discussed, since they are interrelated. The simplest formulation of the problem, exemplified by the fan-shaped antenna patterns, will be developed and will have the advantage of introducing the simplicity of a single variable. The resolving capabilities of three different antenna patterns, each derived from the same over-all aperture, will be found. Two of these patterns have been realized already with a slotted waveguide array (Covington and Broten 1954; Covington and Broten 1957), while the third remains to be developed.

THE OBSERVED BRIGHTNESS PROFILE OF AN OBJECT

The general equation relating the observed temperature distribution produced by scanning a distant object with a fan-shaped radiation pattern to

[1]Manuscript received July 15, 1959.
Contribution from the Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5381.

the true temperature distribution of this object may be written for the single dimension under consideration as

$$R(\theta) = \int_{-\infty}^{\infty} A(\theta - x) S(x) dx, \tag{1}$$

where $R(\theta)$ is the equivalent temperature of the antenna radiation resistance when the maximum of the antenna pattern is at angular distance θ from some reference point,

$A(\theta - x)$ is the normalized power pattern of the antenna as a function of the angular distance $(\theta - x)$ from the maximum of the antenna pattern,

$S(x)$ is the temperature brightness of the source, as a function of the angle x from the reference point.

This equation shows that the observed brightness profile obtained from scanning the object with the antenna pattern differs from the true one as a result of an averaging process introduced by the antenna pattern itself. An alternative relation between the Fourier transforms in the spatial frequency variable "s" is given by the equation

$$\overline{R(s)} = \overline{A(s)} \cdot \overline{S(s)}, \tag{2}$$

where $\overline{R(s)}$, $\overline{A(s)}$, and $\overline{S(s)}$ are the Fourier transforms for the three functions defined for eq. (1); for example,

$$\overline{R(s)} = \int_{-\infty}^{\infty} R(\theta) \exp(-i2\pi s\theta) d\theta. \tag{3}$$

The variable "s" is the number of angular wave crests described per unit angle of θ.

The use of the concept of spatial frequency suggests the analogy that the antenna may be regarded as a filter and that the spectrum of the antenna pattern may be spoken of as the antenna band pass. For an antenna whose aperture is large, as measured in number of wavelengths, the band pass has two features:

(i) it is zero for all values greater than a cutoff frequency s_c given by the relationship

$$s_c = D/\lambda, \tag{4}$$

where D is the length of the basic aperture under consideration, and λ is the wavelength of observation;

(ii) it takes on a definite shape in the region below cutoff, attenuating some frequencies in relation to others.

In the process of making an observation, these two characteristics of the antenna band pass make their imprint on the spectrum of the object and account for all the distortion introduced into the true profile of the object. In terms of frequencies as given by eq. (2), this amounts to the complete removal of all frequencies which are greater than cutoff frequency and, in general, to changes in the amplitude and phase of all the lower frequencies which are accepted by the antenna.

The solution of eqs. (1) or (2) for the source function with a known antenna pattern which has a finite cutoff frequency has been called the principal solution, since its spectrum is the same as the spectrum of the true distribution for all the frequencies below cutoff (Bracewell and Roberts 1954; Arsac 1955). This function may not be a good approximation to the source function, since the higher frequencies which were completely suppressed by the antenna are missing from it. However, it is the Fourier approximation with a finite number of terms and is the best that can be obtained unless there is some other knowledge of the source available. To obtain the principal solution from an observed curve taken with an antenna which attenuates some frequencies more than others, a process of successive approximation must be undertaken that will restore all the frequencies below cutoff to their original values. This restoration process will be limited by the presence of errors, such as receiver noise, since the restoration increases the amplitude of a given frequency regardless of its origin. The principal solution may also be obtained directly when the antenna has a constant band pass from zero to the cutoff frequency. In this case, the antenna power pattern is represented by the function $(\sin 2\pi s_c\theta)/(2\pi s_c\theta)$ and may be approximated very closely in practice (Covington and Broten 1957). No restoration need be done with the observed results and it is in this sense that this particular antenna pattern with its relatively large side lobes is an optimum.

CLASSIFICATION OF ANTENNA PATTERNS BY MEANS OF BAND PASS

The spatial frequency spectrum of the antenna pattern provides an alternative means of describing the directional properties of an antenna and will be used to describe some of the various types of antennas which have been developed within the past few years for radio astronomy. For the purpose of evaluation, the antennas considered are all characterized by the collection of energy from the same aperture, but are physically different in the method of using this aperture. In Fig. 1, some of these configurations are shown together

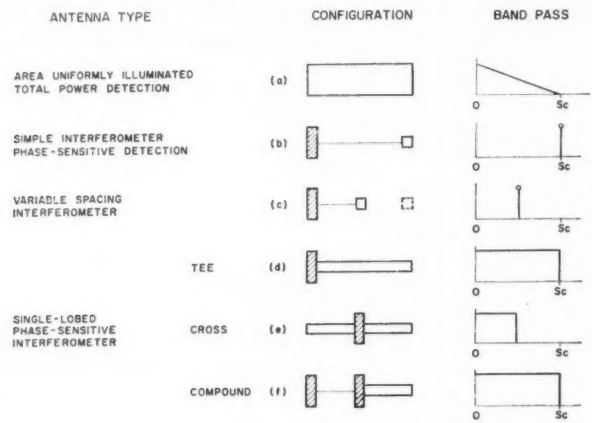


FIG. 1. Various types of antennas, their physical configurations, and the corresponding angular spatial spectra.

with the corresponding band passes. In all cases, with the exception of the cross, the band passes have a common cutoff frequency. In the construction of the successively larger and larger antennas which are required in some radio astronomy studies, it will be readily appreciated that, at some physical size, further fabrication will be possible only through the use of individual elements, each erected on a firm foundation and connected together in some manner such as is shown in the configurations of Fig. 1. The important question of power gain, the methods of making the primary elements, and the directing of the beam will not be considered here.

The basic area available for the collection of energy from the incident wave is shown as a rectangle in Fig. 1(*a*) and with suitable antenna design may be made into a single collecting area. When the voltages contributed by all points in the aperture are of equal intensity, the power delivered to the receiver will depend upon the direction of arrival of the wave as given by the receiving power pattern $P(\theta)$ defined by equation

$$P(\theta) = \frac{\sin^2(\pi s_c \theta)}{(\pi s_c \theta)^2}, \tag{5}$$

where the variables have been defined previously. The corresponding band pass in the spatial frequency spectrum, Fig. 1(*a*), is a triangular function with maximum amplitude at zero frequency and minimum at the cutoff frequency s_c. This cutoff frequency is the maximum which may be obtained from the given length of the basic area, when D/λ is greater than unity, and is related to D by eq. (4). All the other antenna patterns which may be obtained by placing different emphasis on the manner of collecting voltages from various points along the aperture have band-pass spectra which are contained within this frequency.

The collecting areas shown in Figs. 1(*b*), (*c*), (*d*), (*e*), and (*f*) belong to various types of phase-sensitive interferometers, all having the same antenna pattern in the short dimension but with various types of patterns in the long dimension. Again uniform collection of voltages from the incident wave is assumed for each of the antenna elements. In the phase-sensitive interferometer, the resultant power pattern is the product of the voltage patterns of the two interferometer elements multiplied by the usual interference term which is related to the distance between the phase centers and is obtained through an electrical process. This is based upon a slow, continuous shifting of the phase angle between the two radio-frequency voltages picked up by the interferometer elements so that these voltages are successively in phase and then out of phase. This leads to the formation of two antenna patterns in the receiver output, one associated with the mean power and the other with the fluctuating power. The desired antenna pattern associated with the fluctuating power is isolated by means of a low-frequency phase-sensitive detector operating synchronously with the slowly rotating radio-frequency phase shifter. In this preliminary discussion, the error arising from the fact that it is necessary to break or shorten one of the slit apertures of the interferometer elements will not be considered, since its effect can be made small and compensated for.

The simple interferometer with the maximum separation permitted by the basic area under consideration, Fig. 1(*b*), has a band pass consisting of a single

line at the cutoff frequency where the band pass of the uniformly illuminated aperture, Fig. 1(*a*), goes to zero. The difference in response at this frequency gives rise to the statement that the interferometer pattern has twice the resolution of the uniformly illuminated array for, although it is multilobed, the lobes are much narrower. This pattern is, of course, only useful when it is restricted by the primary patterns of the interferometer elements to a region known to contain a single source.

In one form of the variable spacing interferometer, Fig. 1(*c*), one of the antennas is mounted on a track so that many stations in the length of the rectangle may be occupied in turn by that interferometer element. These spacings produce antenna patterns having spatial frequencies from zero up to the maximum defined by the simple fixed interferometer, Fig. 1(*b*). In this manner the various spatial frequencies of the source may be observed without distortion and the resultant curves may be combined by appropriate calculations into a single curve representing the profile of the object. The usefulness of this method naturally depends upon the constancy of the source within the interval required to take the many scanning curves.

Configurations 1(*d*), (*e*), and (*f*) make use of the same basic physical space as the others but produce antenna patterns with a single main lobe. In Fig. 1(*d*), the collecting area is occupied by two slit apertures in the form of a T which are operated as a phase-sensitive interferometer. The antenna pattern in this case is given by $P(\theta) = (\sin 2\pi s_c\theta)/(2\pi s_c\theta)$. The spatial band pass is rectangular, all frequencies being accepted uniformly from zero up to cutoff. The advantage over the variable spacing interferometer is that the information may be obtained in a single scan of the object of relatively short duration, thus eliminating the necessity of making subsequent calculations.

When two slit apertures in the form of a cross, Fig. 1(*e*), occupy the basic area, the rectangular feature of the band pass is maintained but the frequency response is reduced by a factor of two. This reduction in resolution arises since the phase centers of the two interferometer elements overlap, and the interference term vanishes completely.

The compound interferometer, Fig. 1(*f*), shows the utilization of the area under consideration by three elementary collecting areas. The antenna pattern in this case is the same as that given by the T and the band pass is uniform with the same cutoff frequency. A physical advantage of this type of construction is that the single long element is only one-half the basic length. In a similar manner the principle of a compound interferometer may be extended to a four-element grating antenna combined with a continuous slit aperture of length equal to the grating gap in order to produce a single-lobed antenna pattern (Barber 1958*a*).

THREE SPECIAL ANTENNA PATTERNS

In the discussion of antennas which use the same length of a given rectangle for their construction two types of band pass with the same cutoff have appeared: one in which the higher spatial frequencies are linearly attenuated to form a triangular band pass, and a second in which all accepted frequencies

appear in equal strength to form a rectangular shape. This second band pass, for reasons previously stated, has been called the optimum for the length under consideration.

A great many other band-pass shapes would appear to be possible, in which the same cutoff frequency is maintained but with various emphasis placed upon the relative strengths of the higher frequencies. A third band pass which follows a cosine variation from maximum response at zero spatial frequency to zero response at cutoff will be considered as providing a shape intermediate between the first two mentioned. This band pass and the corresponding antenna pattern may be conveniently realized with operations based either upon an antenna with an optimum pattern or upon the resultant scan obtained with such an antenna (see Appendix). The first process involves providing for the simultaneous reception with two optimum antenna patterns displaced with respect to each other so that the maximum of one falls upon the first zero of the second (an angular shift equal to one-half of the cutoff period, ϕ_c, to be defined later), while the second may be conveniently realized by the process of smoothing the scanning curve obtained with a single optimum antenna pattern. This involves the shifting of the observed scanning curve by the angle which separates the two antenna patterns and its addition to the original curve. The resultant curve would be equivalent to that obtained with the composite antenna pattern.

The three antenna patterns corresponding to the three types of band pass are shown in Fig. 2. Antenna patterns (2) and (3) would appear to be more desirable than (1) since they have much smaller side lobes. However, the larger side lobes of (1) are related to the uniform band pass and are needed to record the unmodified high-frequency detail of the object being examined. This

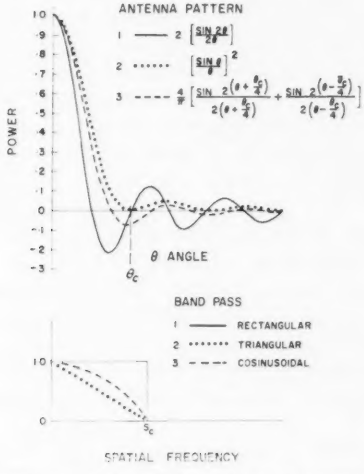


FIG. 2. Three antenna patterns derived from the same aperture and their corresponding spatial spectra.

antenna system with the rectangular band pass enables the Fourier components of a source from zero to cutoff frequency to be received with equal intensity and zero phase shift, and produces scanning or brightness profiles which may be regarded as the Fourier sum approximation with a finite number of spatial frequencies. The use of this pattern thus eliminates the need for further calculation and allows the principal solution to be obtained directly.

SCANNING CURVES OR BRIGHTNESS PROFILES

The relative usefulness of scans obtained by the three receiving patterns without recourse to prior knowledge would vary from object to object. Some appreciation of the relative performance may be readily obtained by reference to theoretical scanning curves or brightness profiles of two models of a small source. The models, a line source of uniform brightness and two equally intense point sources, have been chosen since they represent the simplest of extreme possibilities of variation in source brightness likely to be present in a given small region under study in the sky, as well as providing a means for showing the advantages or disadvantages of the three antenna patterns. By varying the extent of the models—that is, the length of the line object or the separation of the two points—two series of scanning curves have been produced for each of the three antenna patterns. Since the influence of the antenna beam is negligible for uniform sources of large angular extent, the series of curves has been calculated for model angular extents of the order of the width of the main exploring lobe of the antenna. In order to make the curves of general application for use with antennas of different dimensions, a unit angle defined by the cutoff frequency and given the name "cutoff period" has been introduced. This unit, ϕ_c, is given by the relation $\phi_c = 1/s_c = \lambda/D$, and was subdivided into intervals of .048 ϕ_c for purposes of calculation. All the antenna patterns, models, and resultant scanning curves have been tabulated and plotted in these small intervals. Graphical representation of suitable curves for illustration have been made by drawing smooth curves through the plotted points. A selection of the scanning curves obtained with the three antenna patterns for various values of the model parameters are shown for the line source in Fig. 3 and for the two-point source in Fig. 4. These have been normalized so that the central lobe is always equal to unity. The theoretical scanning curves for a parameter of $\phi_c = 0.2$, with either model, differ so slightly from the antenna patterns ($\phi_c = 0$), that for the purpose of visual inspection either may be regarded as the reference antenna pattern.

The small differences existing between the various scans are most readily revealed by measurement of the scan features and contain information which is related to the nature of the source. Although there are many features of the curves which might be used for a description, only the more obvious ones likely to be of practical value in an examination of records will be considered. These are

(1) the half width of the main lobe,
(2) the visibility of the side lobes adjacent to the main lobe,
(3) the visibility of the central dip in the main lobe.

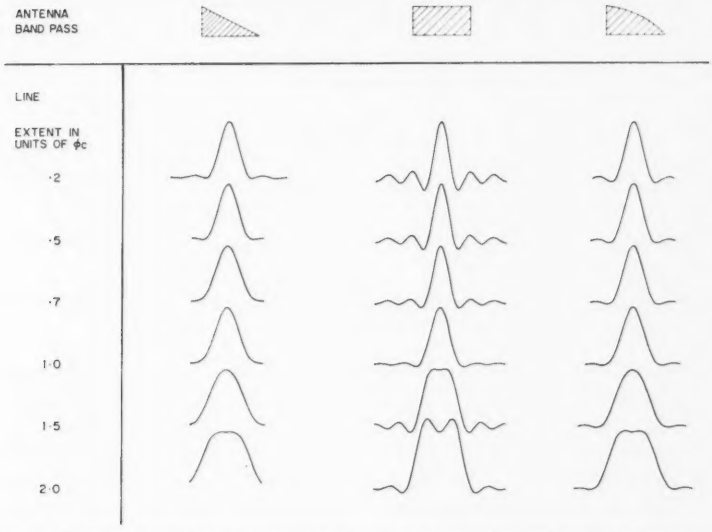


FIG. 3. Calculated profiles of line sources of various extents obtained with the three antenna patterns.

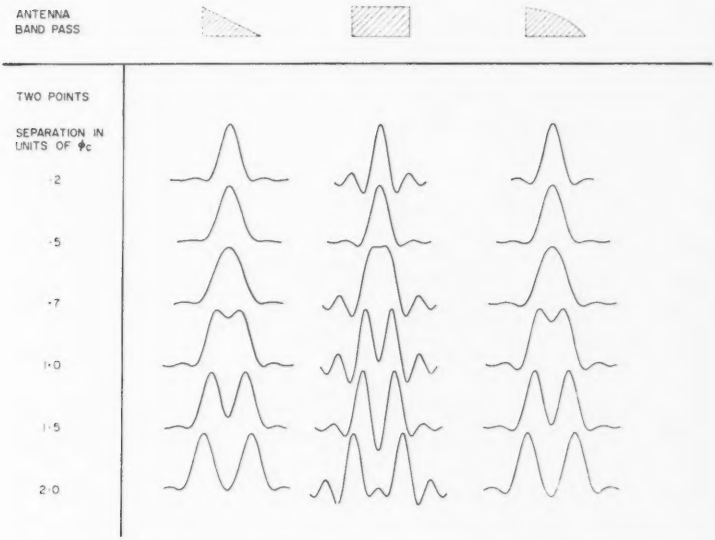


FIG. 4. Calculated profiles of two point sources at various separations obtained with the three antenna patterns.

A graphical representation of a typical brightness profile or scan in the vicinity of the main lobe is shown in Fig. 5, and arrows indicate the dimensions which directly or indirectly describe the scan characteristics. It will be noted

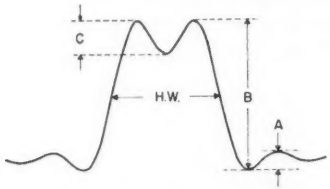


FIG. 5. Definitions of scan or profile characteristics. (1) Half width, H.W.; (2) visibility of first side lobe, A/B; (3) visibility of central dip, C/B.

that no consideration has been given to the zero level and that the strength of the main lobe is taken from the peak to the first minimum regardless of the sign of this minimum value. The half width has been taken at one-half strength of the main lobe. Numerical values for the three listed characteristics, half width, visibility of first side lobe, and visibility of central dip, have been calculated for the scanning curves presented in Figs. 3 and 4 as well as for other curves with intermediate values of the model parameter. Figures 6, 7, and 8 show these values plotted as functions of the line width or the point separation expressed in terms of the cutoff period.

Examination of these functions shows that the response of antenna patterns (2) and (3) as the angular extent of each model is increased is very similar. However, the response in the case of antenna pattern (1) is considerably different. In the case of the two-point model, much more rapid broadening takes place from $\phi_c = 0.5$ to $\phi_c = 1.0$, the central dip in the main lobe appears sooner, and the dip is more rapid; the variation in the height of the first side lobe is much greater. For the line model similar differences are observed

TABLE I

Magnitude of profile characteristics for various angular extents of the two models

		H.W. in ϕ_c			Central dip, C/B			Side lobe A/B		
ϕ_c	Model	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0	2 Point	.69	.85	.89	—	—	—	.283	.048	.093
	Line									
0.5	2 Point	.86	1.01	1.07	—	—	—	.095	.005	.035
	Line	.72	.90	.94	—	—	—	.200	.020	.060
1.0	2 Point	1.77	1.92	1.95	.74	.18	.25	.380	.065	.120
	Line	.93	1.08	1.16	—	—	—	.065	—	.025
1.5	2 Point	2.05	2.32	2.36	1.24	.83	.94	.190	.030	.075
	Line	1.60	1.53	1.60	.03	—	—	.155	—	.030

between the response of antenna pattern (1) and the other two, although these differences do not appear until the angular extent of the model is greater. Table I shows the magnitude of the three characteristics for values of ϕ_c of 0.5, 1.0, and 1.5, derived from the curves of Figs. 6, 7, and 8.

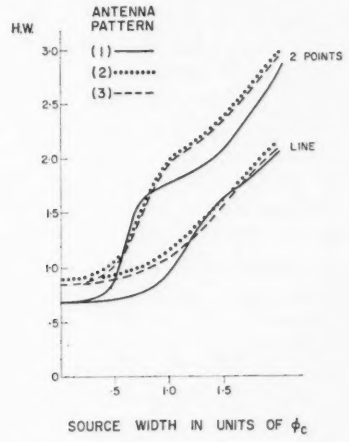


FIG. 6. Plot of the half width of the main lobe of the profile for various angular extents of the two models.

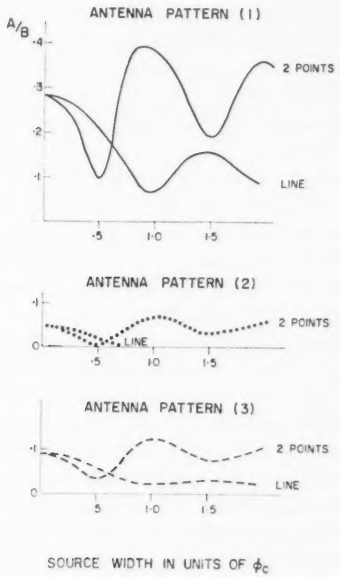


FIG. 7. Plot of visibility of the first side lobe of the profile for various angular extents of the two models.

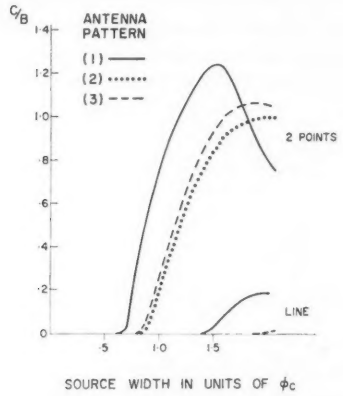


FIG. 8. Plot of visibility of the central dip in the profile for various angular extents of the two models.

RESOLVING POWER

The ability of a given antenna pattern to separate two point sources is expressed most readily in terms of the appearance of the cental dip in the main lobe, and as in the case of the optical telescope, the least discernible angular separation may be used as a measure of instrument performance. Many choices of a value for the minimum visibility are possible and have been arbitrarily singled to one by adopting the criterion suggested by Rayleigh for solving the corresponding problem for the optical telescope. The minimum visibility is that given by two point sources so situated that the maximum of one single profile falls upon the first minimum of the other. Antenna pattern (2) will be recognized as being identical with the diffraction pattern of a slit and the arbitrary minimum visibility C/B occurs at $\phi_c = 1$, and has a value of 0.178. For antenna patterns (1) and (3) this same value of visibility gives the minimum angular separation of two point sources as 0.75 and 0.95 respectively. The resolving power is reciprocally related to this minimum angular separation and for patterns (1), (2), and (3) the power is 1.33, 1.00, and 1.05 respectively. On this basis of comparison, the sharp-lobed antenna pattern is more effective in separating two sources than the other two.

However, in the presence of a third source of unknown strength and position, this antenna with its large side lobes is not an asset since the interaction between the two principal sources may conceal the third one, or the third source will introduce what amounts to an error in determining the various characteristics of the scan. For antenna pattern (1), the strength of any interfering source varies inversely as the angular distance of the source. Antenna patterns (2) and (3), although much broader than (1), do not possess this degree of susceptibility to interference since the interference of a nearby source varies in these cases inversely as the square of the angular separation.

Of these two, pattern (3) with its cosine-shaped band pass is to be preferred since it has the sharper main lobe and the side-lobe level is only slightly increased. However, appropriate smoothing of the observed scan obtained with the optimum pattern will accomplish the same result (Barber 1958*b*).

EQUIVALENT WIDTH OF LINE SOURCE OR SEPARATION OF TWO POINTS

With an antenna whose pattern is one of those considered here, the graphs of the theoretical curve characteristics may be of practical use in determining the nature and size of the source being examined. By a process of measuring the characteristics of an observed scan and making comparison with the theoretical curves, the most suitable model may be chosen. The presence or absence of a central dip is most useful in determining the nature of the source rather than the half width of the main lobe or the strength of the side lobes. Once the nature of the source is determined, the relative broadening of the half width of the main lobe of the profile in terms of the original antenna pattern for various angular extents of the two models will be the most significant, and the curves in Fig. 9 will be of special use. For source widths less than $0.5\phi_c$ the broadening of the scanning curves is so gradual as to be of little practical value in determining the extent of the source.

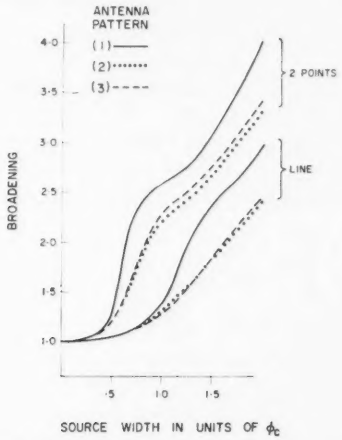


FIG. 9. Plot of the relative broadening of the half width of the main lobe of the profile for various angular extents of the two models.

There are various complicating factors that must be considered: sources are not often as simple as the two models chosen for study here; the true antenna pattern of the observing equipment will have unknown errors; and the presence of receiver noise will cause errors in reading the records. The last may perhaps be readily allowed for by considering the appropriate uncertainty superimposed upon the theoretical curves, thus determining the exactness with which any scan

feature may be read. In spite of these complications the present development for a single dimension does show what may be expected when sources are examined with antennas of the same physical extent but different band-pass characteristics.

REFERENCES

ARSAC, J. 1955. Optica Acta, **2**, 112.
BARBER, N. F. 1958*a*. Proc. IRE, **46**, 1951.
——— 1958*b*. New Zealand J. Sci. **1**, 35.
BRACEWELL, R. N. and ROBERTS, J. A. 1954. Australian J. Phys. **7**, 615.
COVINGTON, A. E. and BROTEN, N. W. 1954. Astrophys. J. **119**, 569.
——— 1957. IRE Trans. on Antennas and Propagation, **AP-5**, 247.

APPENDIX

Methods for obtaining various antenna patterns are suggested by operations which may be performed with equations (1), (2), and (3), while still retaining what might be called a linear relationship between the received power and the source temperature; i.e., limiting the investigations essentially to the single operation of addition.

A direct approach to the realization of this operation involves the construction of two complete antenna feeders so that two distinct antenna patterns are formed at angles $\theta+\tau$ and $\theta-\tau$. The powers R_1 and R_2 delivered to the receivers are given by the equations

$$R_1 = R(\theta+\tau) = \int_{-\infty}^{\infty} A(\theta+\tau-x)S(x)dx, \tag{6a}$$

$$R_2 = R(\theta-\tau) = \int_{-\infty}^{\infty} A(\theta-\tau-x)S(x)dx. \tag{6b}$$

Simultaneous reception of these two powers as a sum, M, is expressed by the equation

$$M(\theta) = \int_{-\infty}^{\infty} [A(\theta+\tau-x)+A(\theta-\tau-x)]S(x)dx. \tag{7}$$

In this equation, a new antenna pattern appears to be formed as the sum of two basic patterns which are separated by an angle 2τ.

The possibilities of a choice of this shift angle are perhaps better revealed through repeating the analysis in the spatial frequency domain. The individual responses to the two antennas are given by the equations

$$\overline{R_1(s)} = \overline{A_1(s)} \,.\, \overline{S(s)}\,, \tag{8a}$$

$$\overline{R_2(s)} = \overline{A_2(s)} \,.\, \overline{S(s)} \tag{8b}$$

where, for example,

$$\overline{A_1(s)} = \int_{-\infty}^{\infty} A(\theta)e^{-i2\pi s(\theta+\tau)}d\theta = e^{-i2\pi s\tau}\,\overline{A(s)}, \tag{9a}$$

$$\overline{A_2(s)} = \int_{-\infty}^{\infty} A(\theta)e^{-i2\pi s(\theta-\tau)}d\theta = e^{i2\pi s\tau}\overline{A(s)}. \tag{9b}$$

Simultaneous reception of the two spectra given in eqs. (8*a*) and (8*b*) is expressed as a sum $\overline{M}$ in the following equation:

$$\overline{M(s)} = [\overline{A_1(s)} + \overline{A_2(s)}] \cdot \overline{S(s)} . \tag{10}$$

With substitution from eqs. (9*a*) and (9*b*), eq. (10) becomes

$$\begin{aligned} \overline{M(s)} &= (e^{i2\pi s\tau} + e^{-i2\pi s\tau}) \overline{A(s)} \cdot \overline{S(s)} \\ &= 2 \cos (2\pi_s \tau) \cdot \overline{A(s)} \cdot \overline{S(s)} . \end{aligned} \tag{11}$$

Thus the effect of adding the responses of two shifted antenna patterns appears in the spatial frequency spectrum as a cosine multiplying factor for the original band pass, with the phase term containing the value of the shift. If 2τ takes on the value of one-half of the cutoff period, i.e., $2\tau = 1/(2s_c)$ and $\overline{A(s)}$ is a uniform spectrum, then the cosine term gives the desired cosinusoidal band pass of the form

$$\cos \frac{\pi}{2} \cdot \frac{s}{s_c} .$$

Using a single antenna, a similar result may be derived indirectly. The response recorded by this antenna may be used to generate a second response which differs from it by the appropriate value by simply considering the recorded response as a function of a variable θ and re-recording it as a function of a new variable $\theta_1 = \theta + 2\tau$. This follows from eqs. (6*a*) and (6*b*) since a shift in angle appears in both sides of the equation. Thus the addition of two separate responses from two antenna patterns is replaced by the addition of two scanning curves, one derived from the other. In practice, this addition could be readily achieved for a fixed scanning rate by using a time-delay circuit to derive the delayed scanning curve for its addition to the original.

THE TEMPERATURE–TIME DEPENDENCE OF THE TRIPLE POINT OF WATER[1]

R. J. Berry

ABSTRACT

The important, but elusive, temperature–time dependency of the triple point of water has been thoroughly investigated in 10 triple point cells from two sources. During the first 2 days after preparation of the cells, the temperature was found to increase by amounts ranging from 0 to 5×10^{-4}° C with the average rise being 2×10^{-4}° C. After the second day the temperature continued to rise at a rate of about 0.1×10^{-4}° C per day for about a week and finally stabilized. In practice, if an ice mantle in a cell is allowed to age for about three days before the cell is used the temperature should be reproducible to about 10^{-4}° C.

A series of experiments are described which suggest that this initial temperature rise may well be due to the growth of crystals in the ice and/or strains in the freshly prepared ice. The slow rise after the second day could be accounted for by crystal growth. These two possibilities are discussed in detail and a formula relating the temperature to crystal size is compared with the observed results.

Tests in pyrex cells up to 5 years old showed that they contain no significant amount of impurities and, therefore, that the segregation of impurities during the freezing process is not likely to be the cause of the initial temperature variations.

On the assumption that the above explanations are true, a number of methods of eliminating this troublesome initial temperature rise were tested. Since none of these tests was completely successful, methods of extending the usefulness of old mantles were examined.

Different methods of preparing and using the cells were critically examined; the earlier method of supercooled freezing was found to be quite inadequate. The effect of different thermal bonds in the thermometer well and of different cell environments was investigated. As a result of this work a new importance is attached to the standard practice of melting the inner layer of ice next to the thermometer well.

The effect of the temperature–time dependency on previous measurements of the difference in temperature between the ice and triple points of water is discussed.

INTRODUCTION

As a fixed reference temperature, the temperature of equilibrium between ice, liquid water, and water vapor (i.e. the triple point) has a most important role in the definition of the two main temperature scales. In 1954 the Tenth General Conference of Weights and Measures (Comité International des Poids et Mesures 1955) adopted an absolute thermodynamic (Kelvin) scale defined in terms of only one fixed point, the triple point of water, to which a value of 273.16° K was assigned. In the International Temperature Scale of 1948 (Comité International des Poids et Mesures 1948) it is recommended that the zero on the scale be realized with the triple point of water (+0.0100° C) rather than with the standard ice point (0° C). In the range 0 to 630° C the international temperature, t, is computed from the resistance, R_t, of a standard platinum resistance thermometer by the equation

$$R_t/R_0 = 1+At+Bt^2$$

[1]Manuscript received June 29, 1959.
Contribution from the Division of Applied Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5384.

where R_0 is the resistance at the ice point and the constants A and B are determined by calibration at the steam and sulphur points. Since it is recommended that R_0 be computed from R_{TP}, the resistance at the triple point of water, the reproducibility of R_{TP} will have a bearing on every determination of temperature in this range.

The triple point of water was first suggested as a substitute for the ice point in 1927 by Michels and Coeterier, who found the triple point reproducible to a few ten thousandths of a degree. An excellent summary of the experimental work that followed on both the ice and triple points has been given by Thomas (1941). During this period the ice point was found to be reproducible to 10^{-4}° C by Thomas (1934) and White (1934) and to 0.5×10^{-4}° C by Beattie (1937). At the same time the triple point was found to be reproducible to 10^{-4}° C by Thomas (1934) and to 0.5×10^{-4}° C by Moser (1929) and Beattie (1937). While it appeared that both points could be realized with about the same accuracy, the triple point was considered to be susceptible to fewer errors because it is independent of both pressure and composition of the atmosphere.

The triple point of water has been realized in glass cells of which our cell shown in Fig. 1 is a typical example. These cells are filled with gas-free distilled water and sealed under vacuum; an axial re-entrant well is provided for the thermometer and is usually partially filled with either water or mercury to improve the thermal contact. Details of the construction of such cells have been given elsewhere by Barber (1954).

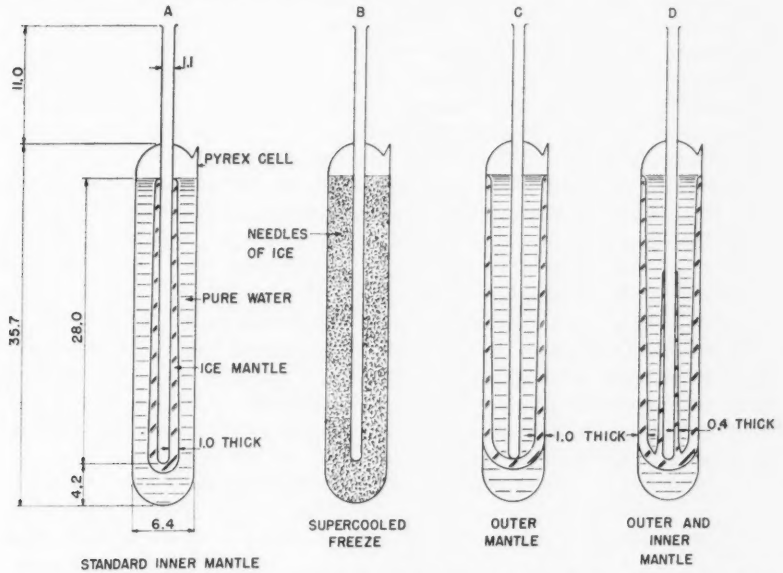


FIG. 1. This shows the construction of the triple point cells and the four different methods of ice formation used. All dimensions are given in centimeters.

Early investigators such as Michels and Coeterier (1927), Moser (1929), and Thomas (1934) prepared their cells by placing them in a brine bath and supercooling the water by 5 to 10° until a spontaneous freezing took place, filling the cell with radially elongated needles of ice. A disadvantage of this method was pointed out by White (1934), who found that after a few hours the ice tended to float up leaving no ice–water interface near the bottom of the well; this resulted in an unsteady temperature reading. By freezing an inner mantle of ice around the thermometer well, using solid carbon dioxide in the well as a refrigerant (Fig. 1A), White found he could keep a cell usable for periods up to 1 week. In addition he pointed out that this new method of freezing should concentrate the impurities in the water, since they are less soluble in ice than in water, and then when a thin layer of ice is melted next to the well a pure ice–water interface would be obtained. White found, however, that the concentration and diffusion of impurities in the water was exceedingly troublesome and that the inner melting of pure ice gave only temporary relief. Nevertheless this inner mantle technique was later used quite successfully by Beattie (1937) and Stimson (1945) and has now become standard practice.

Recent investigations with the triple point cells have confirmed the high reproducibility of their equilibrium temperature, but have also discovered some limitations. Using a resistance thermometer, Barber (1954), Stimson (1955), McLaren (1957), and the author (Berry 1958) found that the temperature of different triple point cells agreed to within 10^{-4}° C. Barber and Berry showed that pyrex cells up to 4 and 5 years old still gave the same temperature as new cells within 10^{-4}° C. Because the deuterium oxide content of water could vary from place to place and with the distillation techniques used to purify the water, Barber felt that the isotopic composition must be considered in order to get all triple point cells to agree to better than 10^{-4}° C. Both Barber and McLaren found that the triple point temperature of some cells was low by about 10^{-4}° C on the day that the mantle was frozen. While Barber attributed this effect to instrumental limitation he also suggested that it could be due to strains in the freshly prepared ice. Recently Stimson (1958) has reported that some of his earlier work gave statistical evidence that the temperature of 3-day-old ice was about 10^{-4}° C higher than 0-day-old ice. In this laboratory we found that each time a new mantle was frozen in a cell the temperature usually rose 1 to 3×10^{-4}° C in the first 2 days* (see Fig. 2) after which it remained constant to about 10^{-5}° C for periods of up to 4 weeks. In obtaining this high reproducibility of the equilibrium temperature we also demonstrated that sudden changes in R_{TP} due to accidental straining of the platinum wire in the thermometer could be avoided by careful handling.

Because the triple point of water is the most important temperature standard and because the initial temperature variations were a factor of 10 greater than the reproducibility of the equilibrium temperature, it was felt to be worth while to try to eliminate this initial temperature variation. The first experi-

*Measurements were usually begun 1 to 2 hours after freezing the mantles.

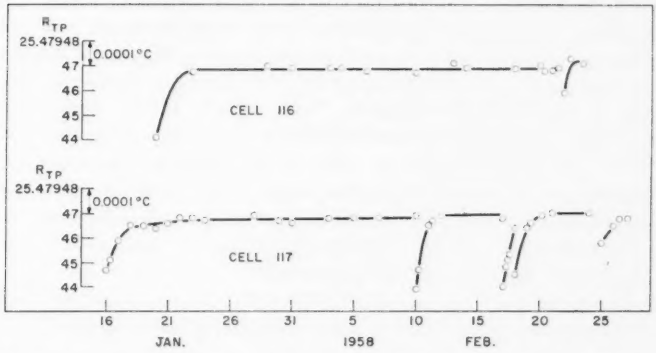


FIG. 2. Variations of R_{TP} measured in two triple point cells with a Meyers thermometer (from Berry 1958). The first measurement shown with each new mantle was taken about one hour after it was frozen.

ments, which were designed to accentuate impurity or strain effects, led to an understanding of the reasons for the initial temperature rise; a number of unsuccessful attempts to eliminate the rise then followed. During the course of this work different methods of preparing and using the cells were critically examined.

EQUIPMENT AND PROCEDURE

The resistance measurements were made with a Leeds and Northrup high-precision G3 Mueller bridge using a type HS galvanometer as the detector. Use of an effective optical path of 9 meters and current reversals allowed interpolation between steps of the 10^{-5}-ohm dial to $\pm 2\ \mu\Omega$ for a bridge setting of 0 to 40 ohms and a thermometer current of 1 ma. At the thermometer sensitivity of 0.1 ohm per degree it was thus possible to measure temperatures with a precision of 2×10^{-5}° C. The bridge was given periodic internal calibrations and interpolated bridge corrections were applied to the daily measurements. While this correction compensated for any drift in the bridge coils relative to the internal reference resistor, X (i.e. the sum of the 10 1-ohm coils) the value of R_{TP} was still affected by any change in X itself. The resistance of X, therefore, was determined periodically in terms of the ohm as maintained in Canada and found to change less than 10 $\mu\Omega$ (1 p.p.m.) over a period of 2 years. Since the ohm as maintained in Canada drifts less than 1 p.p.m. a year (by comparison with the value presently accepted by the B.I.P.M. at Sèvres), the value of X probably does not change any more than a few microhms per month. This estimate appears to be confirmed by the fact that the previous triple point measurements (Berry 1958) showed no appreciable drift in the equilibrium value of R_{TP} over a period of 4 weeks. More details about the bridge and its calibration may be found in the above reference.

Most resistance measurements consisted of 12 observations, current (1 or

2 ma) and commutator (normal or reverse) sequence being 1N 1R 1R 1N, 2N 2R 2R 2N, 1N 1R 1R 1N; zero current resistances were obtained by extrapolation from the measurements at 1 and 2 ma. Whenever less precision could be tolerated, for example during the initial temperature rise in the cells, the above sequence was shortened to 1N 1R, 2R 2N, 1R 1N. Before and after each measurement the residual resistance of the bridge when the dials were set to zero was determined and its mean value subtracted from the resistance readings. All the measurements were made with one Meyers, single layer helix, platinum resistance thermometer (S163) which had been annealed at 444° C. The thermometer was always placed in the triple point cell for at least 15 minutes before readings were begun.

Ten cells were used for this work; the age, manufacturer, and difference in temperature between cells are given in Table I. Unless otherwise specified,

TABLE I

Description of the triple point cells used in this work

Cell No.	Maker	Date manufactured	R_{TP} relative to cell 27	
			Feb. 1958, $\mu\Omega$	Feb.–Apr. 1959, $\mu\Omega$
27	N.R.C.*	Apr. 1954	0	0
21	N.R.C.*	Apr. 1954	+1	
24	N.R.C.*	Apr. 1954	+4	
29	N.R.C.*	Apr. 1954	+1	
25	N.R.C.*	Apr. 1954		−9‡
116	J. & J. Co.†	Aug. 1955	−6	
117	J. & J. Co.†	Aug. 1955	−6	0
201	J. & J. Co.†	Feb. 1958	+3	+9
202	J. & J. Co.†	Feb. 1958	+6	
206	J. & J. Co.†	June 1958		+14

NOTE: A change in R_{TP} of 10 $\mu\Omega$ is equivalent to 10^{-4}° C.
*The N.R.C. cells were prepared in this laboratory by Dr. R. S. Turgel.
†The J. & J. Instrument Co., Silver Springs, Md., U.S.A.
‡A significant amount of air has leaked into cell 25.

cells were prepared by freezing an inner ice mantle around the thermometer well using solid carbon dioxide (dry ice) and alcohol. The sequence of events was as follows: the first piece of dry ice dropped into the practically dry well caused dendritic ice crystals to form on the outside of the well at the bottom, these dendrites then turned into a solid glass-like mantle which grew outwards and upwards as more alcohol and dry ice were added. The completed mantles were tapered, generally being about 8–10 mm thick at the bottom and about 3–5 mm thick near the surface of the water. It was found that mantles less than 5 mm thick at the bottom could not be used for more than a day or two because a hole would form where the ice touched the bottom of the well. As soon as the mantle was completed the inner layer of ice was melted either by putting room temperature water in the well or by putting chilled water and then a warm glass rod in the well. The water was then left in the well to increase the heat transfer to the thermometer. Before each measurement the cell was given a sharp rotation to check that the

mantle was free-turning on the well and, therefore, that the inner layer of water had not frozen. The mantles were maintained in good condition for periods of up to two months by keeping the cells continuously immersed in a bath of crushed ice.

RESULTS

1. Test for the Concentration of Impurities

The first line of investigation was to determine the effect on the initial temperature rise produced by mixing the inner and outer water and melting the ice back from the thermometer well. The mixing was accomplished by inverting the cell several times so that as the mantle floated up and down along the well it pumped the water into and out of the cavity between the mantle and well. To melt the ice back from the well room temperature water was repeatedly poured into the well.

The results of the above tests on six different mantles in cell 117 are presented in Fig. 3; the times when mixing or melting took place are indicated.

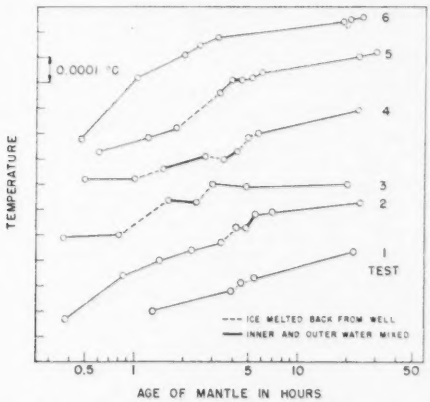


FIG. 3. The effect on the initial temperature variations in cell 117 resulting from mixing the inner and outer water and melting the ice back from the thermometer well. The ice was melted back about 1 mm each time this operation is indicated. The six tests have been arbitrarily moved along the temperature axis for clarity.

Throughout this paper the age of a mantle is taken from the time the mantle was completed. In tests 5 and 6 the mantles were made only three-quarters and half their normal length respectively so as to have a freer interchange between inner and outer water. Since both of these short mantles gave normal equilibrium values for R_{TP}, all the measurements made with them should be reliable. Some idea of the rate of temperature change in an undisturbed cell may be found in the first parts of tests 2 and 6.

In each test a smooth curve, with about the same rate of increase as found in undisturbed cells, could be drawn through the points so that most of them fall on the curve within their experimental error of $\pm 2 \times 10^{-5}$° C. Even if a

few of the sharp changes in the curves are real, it is evident that neither mixing the water nor melting the ice back by as much as 4 mm brought the cell immediately into equilibrium. This shows that the inner and outer water do not contain substantially different amounts of impurities (including deuterium oxide)* and that the initial rise is not due to some property of the innermost layer of ice frozen on the well. It appears, therefore, that the initial rise is a bulk property of the ice.

In order to be absolutely certain that melting the ice back from the well and using short mantles did not alter the final equilibrium value of R_{TP}, the tests described in Table II were made on standard mantles that had already reached equilibrium. Short mantles were obtained by melting those of normal length down from the surface of the water. In the results, also presented in Table II, the only resistance change that is significantly greater than the

TABLE II

Changes in the equilibrium value of R_{TP} caused by melting the ice mantle away from the well and down from the surface of the water. All the measurements in each test were made on the same day

Test	Cell No.	Age of mantle in days	ΔR_{TP}, $\mu\Omega$	Remarks
1	117	35	0	Mantle still giving correct R_{TP} in spite of age. Through usage top 7 cm of mantle had melted and ice was back from well 2 mm
			0	Ice was purposely melted at top until only bottom 7 cm of mantle was left
2	117	5	0	Mantle in good condition
			+2	Ice was melted back 2 mm from well
3	117	5	0	Mantle in good condition
			+8	Ice was melted back 2 mm from well. Reading taken 45 minutes after previous one
			+3	Mantle unchanged. Reading taken 5 hours after first one
4	202	55	0	Mantle still in good condition because it had little usage
			+1	Ice melted back 1 mm
			−3	Inner and outer water mixed
5	201	8	0	Through usage top 2 cm of mantle melted and ice back from well 1 mm
			0	Ice was melted back from well 3 mm
			0	Top 4 cm of ice was melted

experimental error is the value of $+8\,\mu\Omega$ in test 3 ($10\,\mu\Omega$ is equivalent to 10^{-4}° C); a subsequent reading, however, showed that the temperature rise was only temporary. It is evident that even badly deteriorated mantles should still give the correct R_{TP} so long as there is a continuous sheath of ice in the vicinity of the thermometer coil and also that heat conduction down the thermometer well is not a serious problem.

2. Effect of Changing the Freezing Rate

The next line of investigation was to try freezing inner mantles at different rates using radically different freezing techniques. Details of the four techniques used are given in Table III while the graphs of the initial temperature

*The possibility of water undergoing isotopic segregation on freezing was first suggested to me by McLaren (1958). He calculated, from data presented by Teis and Florensky (1941), that the initial ice formed might give a temperature about 3×10^{-4}° C low.

TABLE III

Details of the freezing procedure for the tests shown in Fig. 4

Test	Cell	Thermal bond in well during freezing	Refrigerant	Time to freeze mantle, minutes	Thickness of mantle,* mm
1	117	Air	Dry ice	40	10
2	117	Alcohol	Liquid nitrogen (forced into inner tube)	8	8
3	117	Alcohol	Dry ice (seeded at top)	20	6
4	201	Alcohol	Brass rod cooled in liquid nitrogen inserted 5 times	10	4

*The thickness given is that at the level of the thermometer coil.

variations are shown in Fig. 4. In test 2 the liquid nitrogen was forced down the inner of two concentric tubes placed in the thermometer well. In test 3

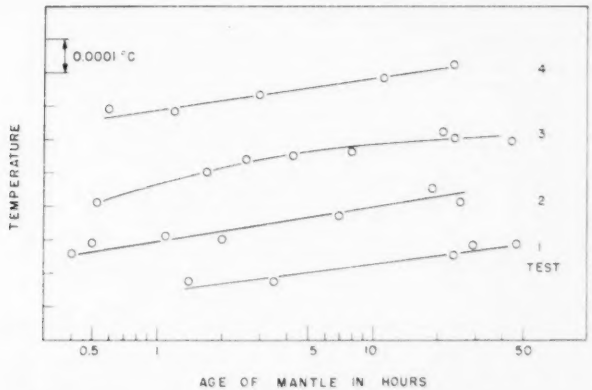


FIG. 4. The initial temperature variations for cells with mantles frozen at varying rates, cf. Table III for details on the freezing techniques. The curves are arbitrarily placed along the temperature axis.

the mantle was started at the top using a piece of ice as a seed and then grown down the well. Since all of the curves in Fig. 4 are similar it appears that changing the radial rate of freezing from 0.2 to 1 mm per minute or eliminating the initial formation of dendritic ice near the thermometer coil has no pronounced effect on the initial rise.

In order to obtain ice that had been frozen still faster it was decided to try the old method of supercooled freezes. For this test cells 27, 117, and 201 were chosen because they had different ages, a spread in their triple point temperatures of $10^{-4°}$ C, and were from two sources. See Table I for details. The cells were supercooled in brine until a spontaneous freezing took place completely filling them with needles of ice. They were then placed in an ice bath and measurements were begun immediately.

After 5 hours it was found that a considerable amount of melting had taken place and that in order to preserve the ice, especially at the bottom, it was necessary to put the cells back in the brine for about ten minutes. This resulted in the ice needles growing into a solid outer mantle about 1 cm thick (Fig. 1C). As the loose center ice needles continued to melt the outer mantle rose until after 10 hours it was touching the bottom of the well. Finally, about twenty hours from the time of freezing, it was necessary to measure the temperature of the outer mantle because the center ice was completely gone in the vicinity of the thermometer coil. With the ice–water interface so far away from the well it took the thermometer about one hour to reach temperature equilibrium after being inserted in the well and an additional 15 minutes after each current change. The change in resistance due to a 1-ma measuring current was increased from 48 $\mu\Omega$ for an inner mantle to 63 $\mu\Omega$ for the outer mantle.

Figure 5 shows the temperature variation in the cells during the first 3 days. It can be seen that all three cells gave practically identical results,

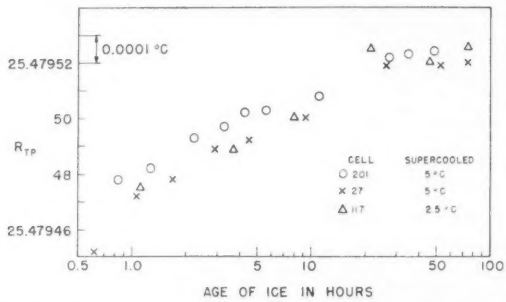


FIG. 5. The initial temperature variations in three different cells with supercooled spontaneous freezes.

starting out about 5×10^{-4}° C low and rising to equilibrium in about one day. Most important of all was the fact that the equilibrium temperatures agreed with each other and with the values for inner mantles to better than 10^{-4}° C. Apparently replacing the cells in brine after they were aged 5 hours did not appreciably disturb the curves. Although the measurements with the outer mantles were extremely tedious to make, their values were highly reproducible. This pronounced depression of the triple point temperature during the first day probably explains why Stimson (1955) found supercooled freezes to be low by 2×10^{-4}° C and why they are generally considered less reproducible than inner mantles.

3. Test for the Amount of Impurities in the Cells

The technique for using outer mantles, once learned, now gave an excellent check of the purity of the water; for if the mantle were frozen slowly the resultant concentration of any impurities near the well would alter the temperature–time curve and possibly the equilibrium value of R_{TP}. The

direction of the changes would depend on whether ordinary impurities or deuterium oxide had been concentrated. Employing the same three cells, outer mantles were frozen at a rate of 0.2 mm per minute by placing the cells in a brine bath ($-10°$ C) and using an ice seed to prevent supercooling. After about 1-cm thickness of ice was frozen against the outer wall the cell was warmed until the mantle floated up and touched the bottom of the well. The cell was then placed in an ice bath and its temperature followed until equilibrium was reached. As a further check on impurity effects a thin inner mantle was frozen in a cell already containing a thick outer mantle (see Fig. 1D). The layer of water between the inner mantle and the well would almost certainly be mixed with the main body of water and therefore contain a concentrated amount of impurities. This latter procedure allowed the measurements to be performed in less time and with greater precision.

The results of these two tests as well as those for the supercooled freezes and standard inner mantles are all shown in chronological order in Fig. 6.

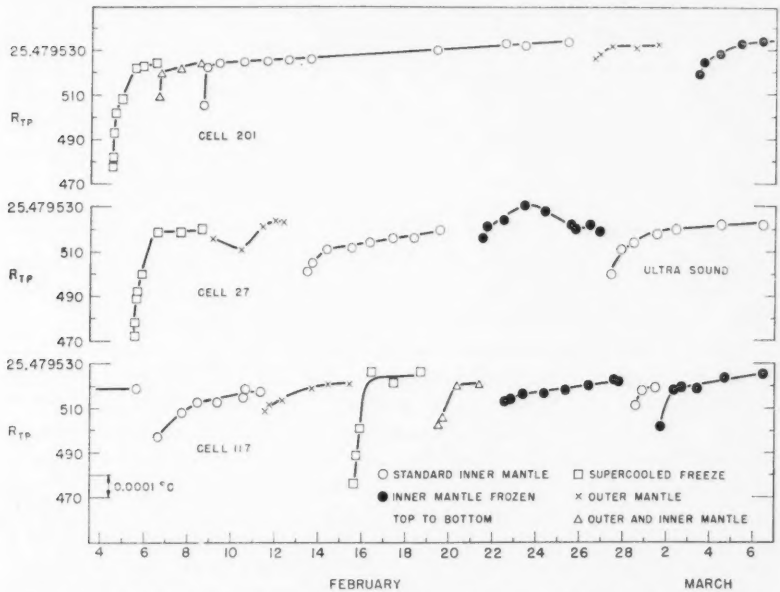


FIG. 6. Variations of R_{TP} measured in three different triple point cells with the four different types of ice formations shown in Fig. 1. The first measurement shown with each new mantle was taken about one hour after it was frozen. The February 5 measurement in cell 117 was taken with a 1-month-old mantle.

Allowing for a slow drift in the measuring equipment, it can be seen that all four types of ice formations in each cell gave the same equilibrium temperature within $\pm 3 \times 10^{-5}$° C. Also both the outer mantles and the inner and outer mantles gave normal temperature–time curves. The absence of any abnormal effect in the tests described above and the fact that old and new

cells from both sources have the same triple point within 10^{-4}° C show rather conclusively that the total amount of impurities in the cells is extremely small. It seems certain, therefore, that the initial temperature variations in our cells, including those for the supercooled freezes, are not due to the concentration and diffusion of impurities. The effect can only be a bulk property of the ice.

4. Evidence of a Slow Rise in R_{TP}

In Fig. 6 most of the results obtained with inner mantles show a slow linear rise of from 0.5 to 1.5 $\mu\Omega$ per day after the comparatively rapid rise during the first day or two. This is most clearly demonstrated by the February 8 mantle in cell 201 which gave a linear rise of 9 $\mu\Omega$ from February 10 to February 25. The resistance drift in the bridge and thermometer as deduced from the values of R_{TP} for cell 117 for mantles 5 days old was about $+7$ $\mu\Omega$ over the month, not nearly enough to account for the observed rise. In previous work with thermometers in the range 0 to 100° C we have detected similar long-term drifts of from 0 to 10 $\mu\Omega$ per month.

At the same time as R_{TP} was rising visible changes were noticed in the crystal structure of the ice. When a mantle was first frozen the crystal boundaries were barely discernible, but after several hours they became quite pronounced. Although the crystals were only a few square millimeters in cross section at first, they slowly increased in size until after a week or two some of them had cross sections of several square centimeters. That the visible crystal boundaries are liquid was demonstrated by putting dry ice down the thermometer well of a cell containing an old mantle; this caused the boundaries to disappear and leave clear ice. However, several hours later, as melting took place, they became visible again revealing the same large crystals that had been present in the mantle before refreezing.

5. Two Explanations for a Rising Temperature

There are two possible explanations for the initial rise in terms of a bulk property of the ice. The first one, proposed by Barber (1954), is based on the annealing of strains in the freshly prepared ice. Since the Clapeyron relation gives $\Delta T/\Delta P = -0.008$° C per atmosphere at the triple point, reasonably small strains could cause the observed lowering of temperature. In the case of the supercooled freezes the rapid freezing should produce greater strains in the ice and therefore a greater temperature drop. It is unlikely, however, that annealing would take longer than a day or two.

The second explanation for the initial rise of the triple point temperature, suggested by Beattie and Stimson (1958), involves the effect of crystal size and shape on the temperature. Since the free energy per mole of small crystals is greater than that of the large crystals because of their greater surface, the vapor pressure of small crystals must be greater and therefore it can be seen from the Clapeyron relation that the melting temperature must be lowered by the fine state of subdivision of the solid. The equation relating the temperature of equilibrium between two phases and the curvature of the interfacial surface (see, for example, Epstein 1937) is $\Delta T = \sigma CT/\rho L$, where, in this particular case,

σ = the interfacial tension between ice and water in ergs cm^{-2},
ρ = the density of the solid in g cm^{-3},
T = the triple point temperature in ° K,
L = the heat of fusion per gram at the triple point,
$C = (1/r_1)+(1/r_2)$ where r_1 and r_2 are the radii of curvature of the interface in two perpendicular planes in centimeters.

Although the crystals in the triple point cells vary considerably in size and shape we shall assume for the purposes of calculation that they are spherical in shape, i.e. $C = 2/r$. Values of σ ranging from 10 to 49 ergs cm^{-2} have been reported (Jacobi 1955) but a value of 20 appears to be currently accepted and is used here. The above equation then reduces to

$$\Delta T = (0.4\times10^{-5})/r.$$

Substituting our observed increase in radius of approximately 1 mm to 5 mm during the first 2 weeks, we obtain an increase in temperature of 3×10^{-5}° C. In order to account for an additional rise of 2×10^{-4}° C in the first day the radius would only have to change from 0.2 to 1 mm during this time. Considering that the actual curvature is much greater than that assumed, particularly at the sharp corners of the crystals, the effect of crystal size and shape is of the right order of magnitude to account for the results. In the case of supercooled freezes the rapid freezing would produce still smaller crystals and therefore cause a greater lowering of the temperature. Because the relief of strains and the growth of crystals are both accomplished by recrystallization it is difficult to consider the two effects independently.

During the course of our investigation, 4 out of some 50 mantles have given temperature–time curves that peaked by about 10^{-4}° C as long as 2 days after preparation. These are shown in Fig. 7, with the one in cell 27 also appearing in Fig. 6. Both new and old cells exhibited this phenomenon. The mantle on cell 27 was frozen from the top of the well down; the others were frozen in the standard way. This anomaly is not readily explained by the above theories alone.

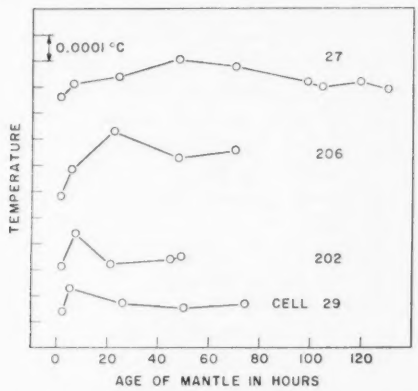


FIG. 7. The four temperature–time curves that showed temperature peaks. The curves are arbitrarily placed along the temperature axis.

6. Attempts to Eliminate the Initial Temperature Change

Based on the above explanations a number of modifications of procedure were tried in an attempt to eliminate the initial temperature variations. Some of the results will be found in Fig. 8 which is a continuation of Fig. 6.

The following three tests proved quite unsuccessful:

(*a*) Freezing mantles from the top of the well down in the hope of growing larger and less strained crystals near the thermometer coil. See Figs. 6 and 8.

(*b*) Freezing the February 27 mantle with cell 27 coupled through oil to a 600-kc ultrasound generator. It was hoped to grow less strained crystals.

(*c*) Freezing the March 7 mantle in cell 117 from the top down using an inner tube to hold the dry ice. It was hoped to distribute the refrigeration more evenly and therefore produce less strained ice.

The next set of tests were partially successful in that they eliminated the large rise during the first day in cell 201 but, unfortunately, only reduced it slightly in cell 117. The results are shown in Fig. 8 for the following tests:

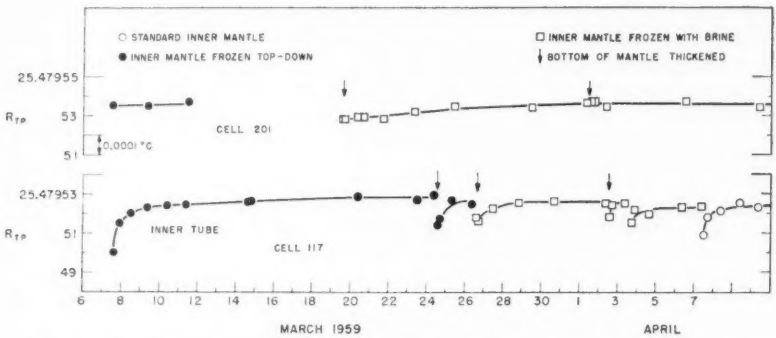


FIG. 8. Variations of R_{TP} measured in two cells. This is a continuation of Fig. 6. The first measurement shown with each new mantle was taken about one hour after it was frozen. The value of R_{TP} on April 14 for the March 19 mantle in cell 201 was 25.479536.

(*d*) Freezing mantles down the well from an ice seed at the top at an extremely slow rate by putting brine (-1 to $-2°$ C) in the well. It took about three hours to build up a 6-mm mantle; even then the very bottom of it remained so thin that after the first measurement it was necessary to thicken the bottom by dropping a small amount of dry ice into the well. See the March 19 mantle in cell 201 and the March 27 mantle in cell 117. It will be noticed that thickening the mantle after the first measurement did not appear to disturb the curve.

(*e*) Freezing the mantle by first building the top half of the mantle with dry ice; then freezing about 1-mm thickness of ice on the bottom half with brine, and finally finishing the mantle with dry ice (see cell 117, April 3). This procedure was considerably less tedious than the one in test (*d*) and gave the same initial lowering of $10^{-4°}$ C.

(*f*) Repairing old mantles by thickening up the very bottom of the mantle

where most of the melting takes place. On April 1 and again on April 14 this was tried on a very old mantle in cell 201 and left R_{TP} unchanged. However, with cell 117 on March 24 and April 2 it resulted in a temporary lowering of 1.5 and 1×10^{-4}° C respectively. Because the new ice formed had the large crystal structure of the old ice, R_{TP} rose to its value for large crystals more rapidly than it would have with a new mantle.

Although our observations of crystal size during the slow rises of R_{TP} illustrated in Fig. 8 showed a direct correlation between size and temperature, none of the preceding methods of freezing a new mantle has proved successful enough in controlling crystal size to warrant the additional effort involved in using them. The graph of the March 19 mantle in cell 201 as well as some of the previous results show that the temperature in a cell will finally stabilize after a week or two. The difference in the behavior of cells 117 and 201 during the last set of tests may be due to the influence, on the strains and crystal size, of traces of impurities or possibly the amount of etching of the glass during the initial cleaning of the cell.

7. Effect on R_{TP} of not Having an Inner Layer of Water

Both Stimson (1955) and White (1934) observed that the temperature of a new inner mantle is quite low before the inner layer of water is formed. They assumed that under these conditions the outer ice–water interface would be controlling the temperature and that the lowering was caused by a concentration of impurities at this outer interface. Since we had found no evidence of impurity concentration, even using the outer mantles, this effect was investigated again using cells 117 and 201.

In order to avoid melting any ice adjacent to the well after freezing the new mantles, the chilled alcohol was left in the well as the thermal bond and the thermometer was precooled to −15° C before placing it in the well. The results presented in Fig. 9 show that the temperature was low by 0.001° C

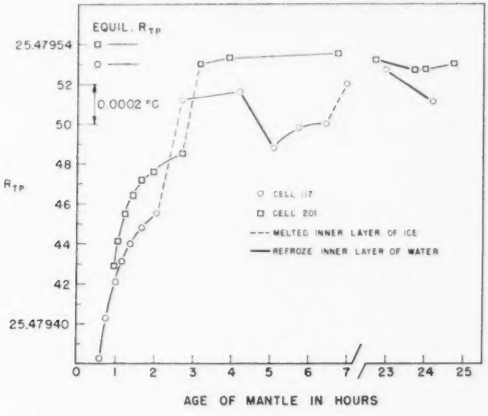


FIG. 9. The effect on R_{TP} of melting and freezing the inner layer of water.

within 1 hour after freezing and rose rapidly during the next hour or two. While the temperature was still low by about 6×10^{-4}° C a thin layer of ice was melted next to the well; this caused the temperature to jump about 5×10^{-4}° C. As a check the inner layer of water was refrozen again by repeatedly inserting a glass rod precooled to −15° C until the mantle was no longer free-turning and the grain boundaries in the ice next to the well had disappeared. In cell 201 this operation left the measured temperature practically unchanged, while in cell 117 it lowered it by about 2×10^{-4}° C.

A more consistent explanation for this large initial rise in temperature is the release of internal stress on the ice through melting and recrystallization. When the inner layer of ice was purposely melted there would then be a sudden release of the pressure exerted by the well on the ice. Refreezing the inner layer of water confirmed previous indications that there is a difference in the behavior of the two cells.

8. Effect on R_{TP} of Changing the Thermal Bonding Medium and the Cell Environment

In order to determine whether the thermal bonding medium in the well has any effect on R_{TP}, a series of tests were made with the following media: water, alcohol, mercury, and aluminum. The aluminum was in the form of a sleeve that fitted closely between the thermometer and the well and was just long enough to cover the platinum coil. In each test the first and last measurements were made with water in the well to guard against changes in the measuring equipment. Unless otherwise specified the height of the liquid in the well was about 28 cm after the thermometer was inserted.

In the results presented in Table IV the ΔR_{TP} shown is the average of two determinations and should be accurate to ±1 μΩ. The first time the thermal

TABLE IV

The effect on R_{TP} of changing the thermal bonding medium in the thermometer well

Thermal bond in well	ΔR_{TP}, μΩ	$\Delta R_{1\,ma}$, μΩ	Thermal conductivity, watts cm^{-1}° C^{-1}
Water	0	48	0.0050
Pure ethyl alcohol	+1.4	69	0.0016
Alcohol + 0.5 ml water (up to 1 hour after mixing)	+16	68	
Water + aluminum sleeve*	+2.7	37	2.1 (Al)
Alcohol + aluminum sleeve	+3.6	43	
Mercury (about 18 cm depth)	−0.8	36	0.063
Water + crushed ice down to within 15 cm of top of thermometer coil	+0.3	47	

NOTE: The values (ΔR_{TP}) shown in the first column are differences between R_{TP} with the thermal bond specified and R_{TP} with water. The values ($\Delta R_{1\,ma}$) shown in the second column are the increases in thermometer resistance produced by a measuring current of 1 ma.

*Dimensions of aluminum sleeve are: length 33 mm, outside diameter 11.1 mm, inside diameter 7.4 mm.

bond was changed from water to alcohol, R_{TP} was found to be high by 15 μΩ up to one hour after the change. Since addition of 0.5 ml of water to pure alcohol produced a similar increase this effect was believed to be due to the

exothermic reaction between water and alcohol. This was finally avoided by flushing the well with alcohol before using it as a thermal bond.

The results show that water, alcohol, and mercury give the same R_{TP} within the experimental error. The increase in R_{TP} of 3 $\mu\Omega$ when an aluminum sleeve is used is not likely to be due to an exothermic reaction between either aluminum and water or aluminum and alcohol. From the last entry in Table IV it appears that if a few pieces of crushed ice should drop into the water in the well accidentally, R_{TP} would not be affected providing, of course, that the ice is not lodged closer than 10 cm to the coil. Although the measuring current heat effect for mercury is slightly less than for water, water is by far the most convenient to use as a thermal bond.

In 1937 Beattie reported that heat leaking downward along the thermometer can be made negligible by immersing the entire triple point cell and part of the thermometer stem in an ice bath. He also found a slight radiation effect of 0.3×10^{-4}° C which was eliminated by wrapping the ice bath in aluminum foil. It was decided to check for these effects by radically changing the cell environment from that normally used in these measurements, i.e. with the cells packed in crushed ice in a stainless steel icebox insulated on the outside with styrofoam.

The results of the tests are shown in Table V. It was found that with the

TABLE V

The effect on R_{TP} of changing the cell environment

Time	R_{TP}	Cell environment
Test 1		
2.10 p.m.	25.479522[7]	In icebox. Crushed ice packed up to top of well
3.00	523[3]	In icebox. Crushed ice only up to water level in cell
4.04	521[5]	In large silvered glass Dewar
4.24	671	At 4.20 cell was taken out of ice and left uncovered in room. Thermometer was left in cell
4.35	675	
4.41	563	At 4.39 cell was completely wrapped in aluminum foil. Still out of ice
4.44	562	
5.20	521[2]	Still wrapped in aluminum foil but back in ice in glass Dewar
5.58	521[8]	In ice box
Test 2		
3.22 p.m.	523[6]	In ordinary galvanized steel bucket blackened on the inside. No insulation on the outside
4.07	523[5]	In stainless steel Dewar*
Test 3—Outer mantle in cell		
1.50 p.m.	532[9]	Cell completely covered with aluminum foil in stainless steel Dewar
2.33	533[1]	In ice box. No aluminum foil

NOTE: All measurements in each test were made on the same day in the same cell. Unless otherwise specified the standard practice of packing the cells in finely crushed ice up to within a few centimeters of the top of the extended thermometer well was followed.

*No difference in R_{TP} with stainless steel Dewar and icebox was ever observed.

ice packed only up to the water level in the cell there was still no effect from heat conduction down the thermometer. Also R_{TP} was not affected measurably by putting the cell in crushed ice in different containers such as the icebox, a glass Dewar, a stainless steel Dewar, and a steel bucket blackened on the

inside or by completely wrapping cells containing inner or outer mantles in aluminum foil and packing them in ice. However, R_{TP} was affected rather seriously if the cell was removed from the ice bath.

9. Test for Air in the Cell

In this laboratory the criterion for a cell containing a negligible amount of air is the audibility of a sharp crack produced by the water hammer when the cell is tipped. With cell 25 this sound is no longer audible and in fact large bubbles of air can be seen in the water by rapidly inverting the cell. Since this cell has a temperature only 10^{-4}° C below the others, the above criterion seems valid. Also of interest is the fact that the air in cell 25 does not significantly alter the temperature–time curve.

DISCUSSION

Two completely experimental determinations of the difference in temperature between the ice and triple points of water have been made by Thomas (1934) and Beattie (1937) and two partly experimental determinations were made by Moser (1929) and Stimson (1945); Stimson obtained a value of 0.0100° C while the others found 0.0098° C. With the exception of Thomas all observers appear to have used their triple point cells on the same day as they were prepared. Both Moser and Thomas used supercooled freezes; Thomas, however, noticed that the temperature difference was lower on the first day and therefore used only measurements made at least one day after the cell was prepared. Stimson appears to have used mantles on the day they were prepared and did not melt the inner layer of ice in order to measure the temperature at the outside of the mantle. It appears then that most of the measurements would be affected by the time dependency of the triple point temperature. However, since the measurements would tend to be slightly low, the values of +0.0100° C finally adopted may be more accurate than previously thought.

From the results presented in Figs. 6 and 8 it can be seen that for a precise intercomparison of triple point cells they should be prepared on the same day, allowed to age at least five days before measurements are begun, and comparisons should be made only between measurements taken on the same day. Even then it is possible that small temperature differences may exist between cells due to the fact that cells of the same age can have different-sized crystals. Since this practice was followed only for the 1959 intercomparison of cells shown in Table I, the 1958 intercomparison shown in the same table is subject to greater error. Although our new cells have consistently higher triple points than our 4- to 5-year-old cells, by about 10^{-4}° C, it is not certain yet whether the old ones have become slightly contaminated or the new ones have a higher deuterium oxide content.

CONCLUSIONS

During the course of this work the temperature–time dependence of some 50 inner mantles was investigated in 10 different pyrex cells from two sources. It was found that in the first day after preparation the temperature increased

by amounts ranging from 0 to 4.6×10^{-4}° C with the average rise being 1.7×10^{-4}° C, while in the second day the temperature increased another 0 to 1.1×10^{-4}° C with the average rise being 0.3×10^{-4}° C. After the second day there was a slow rise of about 0.1×10^{-4}° C per day until finally a week or two later the temperature remained stable within 0.2×10^{-4}° C. This initial temperature rise during the first 2 days is thought to be due to either strains and/or crystal growth in the ice. The long slow rise after the second day is believed to be associated with crystal growth.

A series of tests, which were designed to detect impurities in the water, were made on cells up to 5 years old. The fact that these tests showed nothing unusual and that the old and new cells had triple points that agreed within 10^{-4}° C indicates that the total amount of impurities in our cells is extremely small. For this reason it is extremely unlikely that the initial rise is due to the concentration of impurities ahead of the freezing water. However, it is quite possible that minute traces of impurities have some influence on the temperature by modifying the crystal size, their growth rate, or the strains.

If the water in a cell is contaminated the technique of using an inner mantle and melting an inner layer of water cannot give even temporary relief, because by the time the other factors affecting the temperature have disappeared the inner layer of water would be contaminated. The real value of the inner layer of water is in having a continuous ice–water interface close to the thermometer and the elimination of any pressure that the well might exert on the ice.

The old method of using supercooled freezes was found to result in a temperature that started out about 5×10^{-4}° C low and rose to equilibrium in about one day. The pronounced lowering of the temperature was presumably due to an increase in strains and/or a decrease in crystal size, both of which would occur with rapid freezing. Since the ice was almost completely melted after the first day, this method of preparation is quite inadequate.

In order to make a mantle that is strain free and contains large crystals it is necessary to use an inconveniently slow freezing rate. One way of circumventing this problem is to extend the usefulness of an old mantle by refreezing it with dry ice and building up its deteriorated sections; this may cause some lowering of the temperature but equilibrium is reached more quickly. Another method is to use a mantle for a longer period of time. We have found that they can be used for months provided that there is a continuous sheath of ice in the vicinity of the thermometer coil.

Water was found to be the most convenient thermal bond to use in the thermometer well and gave the same temperature as mercury and alcohol to within 10^{-5}° C. Also it was shown that any effect from heat conduction down the thermometer or from heat radiation was made quite negligible by completely packing the cell in crushed ice.

In order to define the temperature of the triple point of water to better than 10^{-4}° C consideration would have to be given not only to the deuterium oxide content in the water but also to the crystal size. In practice, by allowing a mantle to age about three days before using it the temperature should be reproducible to 10^{-4}° C. Although some of the measurements of the difference

in temperature between the ice and triple points of water were affected by this temperature–time dependence, the value of +0.0100° C finally adopted in the International Temperature Scale (1948) is probably still within the stated limits of 0.0002° C.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the assistance of Mr. N. F. Scardina in all phases of the experimental work. The author also wishes to thank Mr. G. Fischer for many helpful discussions and Drs. H. Preston-Thomas and T. M. Dauphinee for assistance with the manuscript.

REFERENCES

BARBER, C. R., HANDLEY, R., and HERINGTON, E. F. G. 1954. Brit. J. Appl. Phys. **5**, 41.
BEATTIE, J. A., TZU-CHING HUANG, and BENEDICT, M. 1937. Proc. Am. Acad. Arts Sci. **72**, 137.
BEATTIE, J. A. 1958. Private communication.
BERRY, R. J. 1958. Can. J. Phys. **36**, 740.
COMITÉ INTERN. POIDS ET MESURES. 1948. Procès-verbaux des séances de 1948, **21**, T30.
——— 1955. Procès-verbaux des séances de 1954, **24**, T23 and T47.
EPSTEIN, P. S. 1937. Textbook of thermodynamics (John Wiley & Sons, Inc., New York), p. 217.
JACOBI, W. 1955. Z. Naturforsch. **10a**, 322.
MCLAREN, E. H. 1957. Can. J. Phys. **35**, 78.
——— 1958. Private communication.
MICHELS, A. and COETERIER, F. 1927. Proc. Koninkl. Akad. Wetenschap. Amsterdam, **30**, 1017.
MOSER, H. 1929. Ann. Physik (Leipzig), **1** (5), 341.
STIMSON, H. F. 1945. J. Wash. Acad. Sci. **35**, 201.
——— 1955. Temperature. Its Measurement and Control in Sci. and Ind. Vol. **II**, 153.
TEIS, R. V. and FLORENSKY, K. P. 1941. Compt. rend. acad. sci. U. R. S. S. **32**, 199.
THOMAS, J. L. 1934. J. Research NBS, **12**, 323.
——— 1941. Temperature. Its Measurement and Control in Sci. and Ind. Vol. **I**, 159.
WHITE, W. P. 1934. J. Am. Chem. Soc. **56**, 20.

THE PRESSURE-INDUCED ROTATIONAL ABSORPTION SPECTRUM OF HYDROGEN

II. ANALYSIS OF THE ABSORPTION PROFILES[1]

Z. J. KISS[2] AND H. L. WELSH

ABSTRACT

The experimental profiles obtained in Part I for the pressure-induced absorption of H_2 and H_2 – foreign gas mixtures in the region 300–1400 cm^{-1} are separated by a semiempirical method into rotational and translational parts. The pressure-induced rotational lines of H_2 are shown to have a dispersion line shape, modified by the Boltzmann law. The intensities of the rotational parts of the spectra are, on the average, 13% greater than those calculated in Part I on the assumption of quadrupole interaction alone; the residual experimental intensity probably represents the effect of overlap interaction. The intensities of the translational parts of the spectra are of the same order of magnitude as the translational intensities in mixtures of the rare gases. Some additional experimental data on double rotational transitions and on the anomalous density variation of the absorption in H_2–Xe mixtures are given.

INTRODUCTION

In Paper I of this series experimental data on the pressure-induced infrared rotational spectrum of hydrogen were presented and a discussion of the integrated absorption intensities was given (Kiss, Gush, and Welsh 1959). The experimental results comprised measurements on the complete rotational spectrum of hydrogen in the pure gas and in mixtures with nitrogen, helium, neon, argon, krypton, and xenon; total gas pressures up to 250 atmospheres at 300° K were used, and the measurements were extended when possible to 195° K and 85° K. The spectrum, which lies in the region 300–1400 cm^{-1}, showed the expected rotational S lines corresponding to the selection rule $\Delta J = +2$; in general, the lines were very broad with half widths which decreased at lower temperatures. From values of the absorption coefficient, $\alpha(\nu)$, measured across the band, the integrated absorption coefficient, $\int \tilde{\alpha} d\nu = \int [\alpha(\nu)/\nu] d\nu$, was calculated. This coefficient was shown to be of the form $\tilde{\alpha}_1 \rho_{H_2} \rho_f$, where ρ_f is the density of the perturbing gas, except in the case of xenon for which a small cubic term, $\tilde{\alpha}_2 \rho_{H_2} \rho_f^2$, was also necessary. The theoretical value of the binary coefficient, $\tilde{\alpha}_1$, calculated on the basis of quadrupole interaction in collision pairs, was in good agreement with the experimental value only in the case of xenon as the perturbing gas; in all other cases the calculated values were appreciably less than those observed. When the agreement between the experimental and theoretical values was particularly poor, as for helium and neon as perturbing gases, it was noted that the intensity ratio of the S lines did not conform at all with that expected

[1]Manuscript received August 7, 1959.
Contribution from the McLennan Laboratory, University of Toronto, Toronto, Ontario. This research was supported by Grant No. T54 of the National Research Council of Canada.
[2]Holder of a scholarship from the National Research Council, 1956–1957, and of a Canadian Kodak Fellowship, 1957–1958.

from the theory. In these cases the rotational spectrum appeared to be superimposed on a continuum, the intensity of which increased rapidly towards longer wavelengths. It was suggested that this continuum might be a collision-induced *translational* spectrum. This interpretation was confirmed by the observation of similar absorption continua in the same wavelength region in binary mixtures of the rare gases (Kiss and Welsh 1959). Translational continua are presumably produced by the absorption of light by molecules interacting through overlap forces, the energy of the absorbed photon appearing as kinetic energy of the relative motion of the collision partners.

In the further study of the intensity of the rotational spectrum of hydrogen it is necessary to separate the rotational spectrum from the superimposed translational continuum. In the present communication a method of effecting the separation is given, based on the fact that it has been possible to find empirically a satisfactory analytical expression for the shape of an individual rotation line in induced absorption. In addition, some further experimental data are presented; these concern the intensities of double rotational transitions and the anomalous density variation of the intensity of the rotational spectrum in H_2–Xe mixtures.

THE SHAPE OF A SINGLE ROTATION LINE

The first step in the separation of the rotational and translational parts of the absorption profile is to obtain an analytical expression for the shape of a single rotation line. For this purpose it is expedient to use data obtained from low-temperature experiments, since for these the overlapping of the rotation lines is small and the contribution of the translational continuum is less than at room temperature. In Fig. 1 experimental absorption profiles,

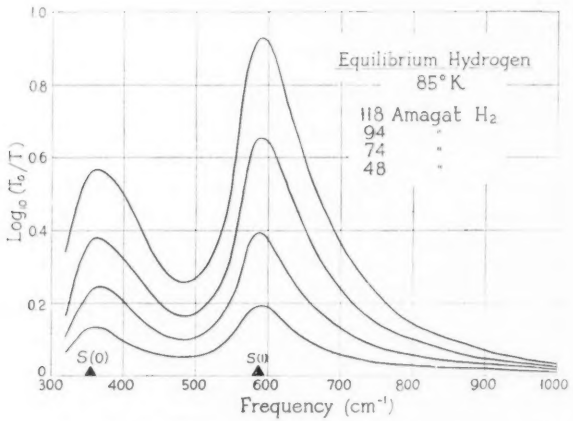


FIG. 1. Absorption of equilibrium hydrogen at 85° K in the region 300–1000 cm^{-1} for various gas densities. The rotational Raman frequencies for the low-pressure gas are indicated on the abscissa axis.

$\log_{10}(T_0/T)$ vs. frequency, are shown for a series of densities of equilibrium hydrogen at 85° K; only the $S(0)$ and $S(1)$ lines are present and the overlapping is not great.

The experimental absorption intensities, $I \equiv \log_{10}(T_0/T)$, contain of course the absorption frequency, ν, which varies greatly over the extent of the rotational band. In the analysis it is convenient to remove this factor and to present the intensities in the form, $\bar{I} = (1/\nu)\log_{10}(T_0/T)$; $\bar{I}$ is thus in effect a transition probability. In Fig. 2 the solid line shows this form of plot for

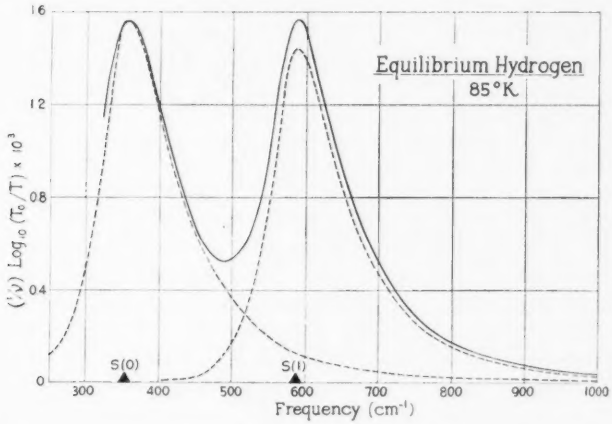


FIG. 2. Separation of the absorption profile for equilibrium hydrogen (85° K, 118 Amagat units of density) into two rotational lines of the same shape.

equilibrium hydrogen at 85° K and 118 Amagat units of density. If it is now assumed that the two lines, $S(0)$ and $S(1)$, have the same shape, a separation can be effected by beginning at the high-frequency side of the band where the overlapping is least; the result is shown by the dotted curves in Fig. 2. The intensity ratio, $S(0):S(1)$, of the two lines is 1.06, as compared with the theoretical value, 1.14, calculated in Part I on the basis of purely quadrupole interaction. A similar separation of the lines for normal hydrogen at 85° K gave the ratio, $S(0):S(1) = 0.56$, as compared with the theoretical value, 0.54. It is therefore concluded that the assumption of similar line shapes is justified.

The profile of an individual rotation line as obtained from the above procedure is asymmetric with the intensity degraded towards higher frequencies; the profile has therefore, at least qualitatively, the appearance of a continuous distribution of summation and difference tones with the rotational frequency. If this interpretation is correct the intensity distribution should satisfy the Boltzmann relationship, $\bar{I}^+/\bar{I}^- = \exp(\Delta\nu/kT)$, where $\bar{I}^+$ is the intensity on the high-frequency side of the line at a point $\Delta\nu$ from the maximum and $\bar{I}^-$

the corresponding intensity on the low-frequency side. In Fig. 3 the plot of $\log_{10}(\tilde{I}^+/\tilde{I}^-)$ vs. $\Delta\nu$ for normal and equilibrium hydrogen at 85° K gives an

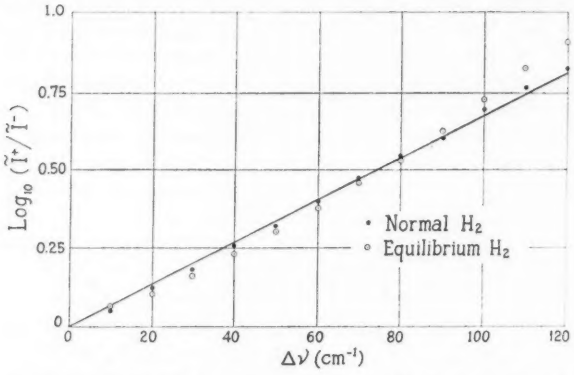


FIG. 3. Graph illustrating the Boltzmann relationship, $\tilde{I}^+/\tilde{I}^- = \exp(\Delta\nu/kT)$, between the wings of the $S(1)$ line at 85° K.

acceptable straight line, the slope of which corresponds to a temperature of 93° K. The asymmetry of the rotational line appears therefore to be expressible by the Boltzmann law.

The most satisfactory analytical expression for the intensity distribution in the high-frequency wing of the line was found to be the dispersion line form,

$$\tilde{I}^+(\nu) = \tilde{I}(\nu_0)/\{[(\nu-\nu_0)/\delta]^2+1\},$$

where $\tilde{I}(\nu_0)$ is the intensity at the frequency ν_0, of the maximum, and δ is the value of $\nu-\nu_0$ for which the intensity is $\tilde{I}_0/2$. In the following, δ will be referred to as the half width of the line. It is evident that the graph of $[\tilde{I}_0/\tilde{I}^+(\nu)]-1$ against $(\nu-\nu_0)^2$ should give a straight line. The plot for the $S(1)$ line of normal hydrogen at 85° K, shown in Fig. 4, indicates that the dispersion form is in

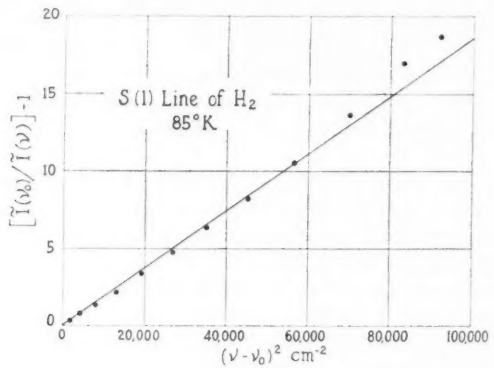


FIG. 4. Graph illustrating the dispersion line form for the high-frequency wing of the $S(1)$ line of normal hydrogen at 85° K.

fact a good representation of the line shape. The small deviation from the straight line in Fig. 4, which occurs at the extreme part of the high-frequency wing [$(\nu_0-\nu) > 250$ cm^{-1}] was observed consistently in all the analyses, and may indicate a real departure from the dispersion shape in this region. The deviation does not, however, constitute a serious limitation on the analysis procedure discussed below.

From the above considerations it is concluded that the shape of an individual rotation line at 85° K can be described fairly accurately by a combination of the dispersion line shape and the Boltzmann law. It was shown by Chisholm, MacDonald, Crawford, and Welsh (1952) that the Q_P and Q_R components of the split Q branch of the induced vibrational band of hydrogen also obey a Boltzmann relationship. The splitting of the Q branch was therefore explained in terms of the participation of the relative kinetic energy of the colliding molecules in the absorption process. In view of the above analysis the wings of the pure rotation lines also seem to be due to an intervention of the translational motion of the molecules in the absorption. There is, however, a notable difference between the two cases: for the Q branch the most probable change in the kinetic energy is not zero, whereas for the rotation line zero change is most probable. This difference between the two cases must eventually be explicable in terms of the two different mechanisms of absorption: overlap interaction in the production of the Q branch absorption, and quadrupole interaction in the case of the rotation lines. The explanation of the dispersion line form for the rotation line must also await a theoretical treatment of the translational energy effect.

ANALYSIS OF THE ABSORPTION PROFILES

The intensity distribution in a single rotation line at 85° K has been shown above to depend on three parameters: $\bar{I}_0$, δ, and the absolute temperature, T. We will now assume that the dispersion line form modified by the Boltzmann law holds for all mixtures and all temperatures with the three parameters suitably chosen. The relative values of I_0 for the different rotation lines will be taken to be those given by the transition probabilities for quadrupole interaction (Kiss *et al.* 1959). Since the temperature is known, there are in effect only two parameters, δ and the total rotational intensity, to be adjusted to fit the experimental data. In the example shown in Fig. 5 the experimental profile for a H_2–Xe mixture at room temperature was accurately reproduced by a variation of the two parameters. This is a particularly favorable case since the translational continuum is very weak and contributes appreciable intensity only in the region of the $S(0)$ line.

A similar analysis was carried out for all the mixtures and temperatures for which experimental data were available. Usually only one pressure was used since the shape of the band changed only very slightly over the pressure range. In the case of H_2–He and H_2–Ne mixtures the separation into components involved some uncertainty since the rotation lines were very broad and the translational continuum relatively intense. In these cases an additional criterion was used in the analysis, namely, that the difference between the synthetic profile of the rotation band and the observed profile should be a

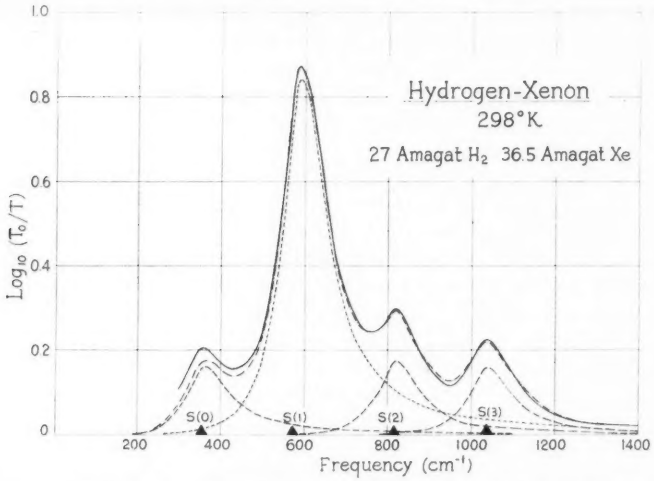


FIG. 5. Construction of a synthetic profile for the enhancement of the rotational absorption of hydrogen by xenon: ——— the experimental profile, – – – the component lines of the synthetic profile, –·–·– the synthetic profile.

continuum decreasing monotonically towards higher frequencies. Figure 6 shows the final result of the separation of the observed profiles into rotational and translational parts for pure hydrogen and H_2 – rare gas mixtures; the curves for the mixtures refer to the *enhancement* of the absorption by the

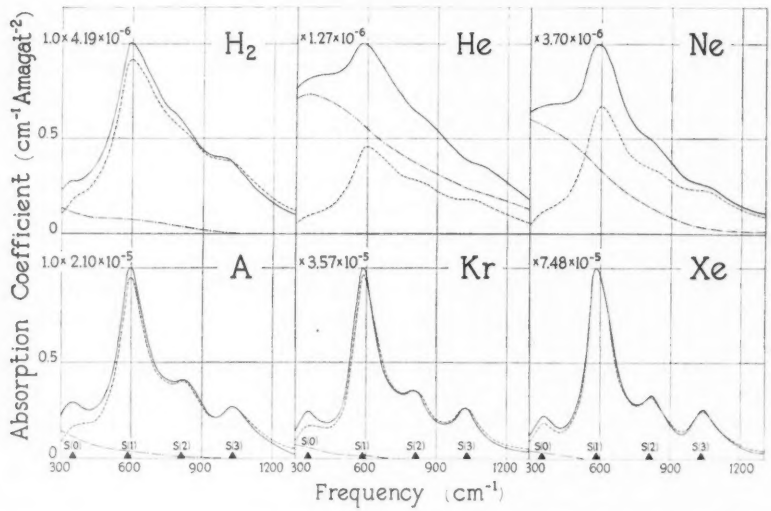


FIG. 6. Separation of the experimental profiles (———) for pure H_2 and H_2 – rare gas mixtures into rotational (– – –) and translational (–·–·–) parts.

foreign gas, i.e. the intensity due to H_2–H_2 collisions has been subtracted from the observed intensities. The ordinate scale used in Fig. 6 is the *specific* absorption coefficient, giving the absorption intensity per centimeter path length per Amagat unit of density of hydrogen per Amagat unit of density of foreign gas.

After the division of the spectrum into rotational and translational parts, it is possible to calculate integrated binary absorption coefficients separately for the two parts. Results for the various experimental conditions are summarized in Table I. For convenient comparison with theory the coefficients

TABLE I

Experimental and calculated values of the absorption coefficients

		Experimental		Calculated	
T, °K	Perturbing gas	α_1 (trans.), $\text{sec}^{-1}\,\text{cm}^6$	α_1 (rot.), $\text{sec}^{-1}\,\text{cm}^6$	α_1 (quad.), $\text{sec}^{-1}\,\text{cm}^6$	δ, cm^{-1}
375	Xe	0.61×10^{-34}	10.9×10^{-34}	10.8×10^{-34}	90
300	n-H_2	0.18×10^{-34}	1.19×10^{-34}	1.06×10^{-34}	155
	N_2	0.45×10^{-34}	3.09×10^{-34}	2.66×10^{-34}	96
	He	0.32×10^{-34}	0.12×10^{-34}	0.108×10^{-34}	145
	Ne	0.442×10^{-34}	0.394×10^{-34}	0.289×10^{-34}	127
	A	0.49×10^{-34}	3.92×10^{-34}	2.94×10^{-34}	94
	Kr	0.25×10^{-34}	5.97×10^{-34}	5.50×10^{-34}	83
	Xe	0.20×10^{-34}	10.9×10^{-34}	11.0×10^{-34}	80
195	n-H_2	—	1.19×10^{-34}	1.06×10^{-34}	124
	N_2	—	3.27×10^{-34}	2.85×10^{-34}	77
	He	0.091×10^{-34}	0.096×10^{-34}	0.099×10^{-34}	120
	Ne	0.12×10^{-34}	0.355×10^{-34}	0.284×10^{-34}	100
	A	—	3.61×10^{-34}	3.17×10^{-34}	75
85	n-H_2	—	1.25×10^{-34}	1.07×10^{-34}	84
	e-H_2	—	1.30×10^{-34}	1.20×10^{-34}	84

are given in the form of transition probabilities, $\tilde{\alpha}_1$(trans.) and $\tilde{\alpha}_1$(rot.), that is, with the absorption frequency factor removed, and in the units, cycles per second per molecule per cubic centimeter of hydrogen per molecule per cubic centimeter of perturbing gas ($\text{sec}^{-1}\,\text{cm}^6$). In column five theoretical values of the rotational binary coefficient, $\tilde{\alpha}_1$(quad.), calculated in Part I under the assumption of quadrupolar interaction alone, are given; the values for pure hydrogen differ slightly from those given in Part I since quantum corrections have been taken into account (Van Kranendonk and Kiss 1959).

The agreement between $\tilde{\alpha}_1$(rot.) and $\tilde{\alpha}_1$(quad.) is now much closer than that between $\tilde{\alpha}_1$(exptl.) and $\tilde{\alpha}_1$(quad.) as discussed in Part I. The removal of the translational part of the absorption has greatly reduced the outstanding discrepancies, as for example in the H_2–He case where the translational absorption is 2.7 times as great as the rotational absorption. However, the values of $\tilde{\alpha}_1$(rot.) are still in most cases greater than $\tilde{\alpha}_1$(quad.), and it is probable that the discrepancy (about 13% on the average) is associated with the neglect of the overlap interaction in the calculation of the theoretical intensities. This question has been considered in detail by Van Kranendonk and Kiss (1959); they show that, although the rotational intensity is mainly

due to quadrupolar interaction, there is a small overlap effect arising for the most part from an interference between the large quadrupolar and the small overlap induced moments. The calculation reproduces very well the observed variation of $\tilde{\alpha}_1$(rot.) with temperature, as well as the change in $\tilde{\alpha}_1$(rot.) in passing from normal to equilibrium hydrogen at 85° K. By using Rosen-type wave functions Van Kranendonk and Kiss were able to calculate the interference effect between the quadrupolar and overlap moments for pure hydrogen. The result was of the correct order of magnitude but of the wrong sign to explain the difference, α_1(rot.) $-\alpha_1$(quad.); this is perhaps an indication of the inadequacy of Rosen-type wave functions in evaluating the angle-dependent part of the overlap moment. Part of the discrepancy between $\tilde{\alpha}_1$(rot.) and $\tilde{\alpha}_1$(quad.) may also be caused by inaccuracies in the calculation of $\tilde{\alpha}_1$(quad.), since some of the constants involved, particularly the intermolecular potential constants for mixtures, may be in error. However, the excellent agreement for H_2–Xe mixtures at 375° K and 300° K is noteworthy.

The values of $\tilde{\alpha}_1$(trans.) obtained from the analysis are of the same order of magnitude ($\sim 0.2 \times 10^{-34}$ sec^{-1} cm^6) as those found for mixtures of the rare gases (Kiss and Welsh 1959). The spectral region (>300 cm^{-1}) to which the observations were necessarily confined contains only a portion of the translational absorption, which presumably has a maximum in the lower-frequency region. Although a quantitative discussion of the values of $\tilde{\alpha}_1$(trans.) is therefore not very meaningful, some points of interest may be noted. The translational intensity in the region > 300 cm^{-1} decreases rapidly as the temperature is lowered as might be expected. The small value of α_1(trans.) for hydrogen must be a consequence of the fact that the overlap interaction in H_2–H_2 collisions is almost independent of the mutual orientation of the molecules so that the induced moment is small; in the extreme case of spherically symmetric-like molecules, as for example in a pure rare gas, the induced moment is zero. In the H_2 – rare gas series α_1(trans.) appears to have a maximum for H_2–A; however, the values for H_2–Kr and H_2–Xe may have a large experimental error since the translational absorption amounts to only a small fraction of the total absorption.

The last column of Table I gives the values of the dispersion half width, δ of the rotation lines, as calculated from the slope ($\equiv 1/\delta^2$) of the straight line in graphs of the type in Fig. 4. The temperature dependence of δ for pure hydrogen and H_2–A is shown in Fig. 7(*a*); the dispersion half width varies as $\sqrt{T}$ and is consequently proportional to the average relative velocity of the colliding pair. The total half width of the line shows a more complicated temperature dependence since, from the above analysis, the low-frequency wing of the line is modified by the Boltzmann law. The variation of the dispersion half width at a given temperature for different collision pairs is illustrated in Fig. 7(*b*), in which δ is plotted against $1/\sigma\sqrt{\mu}$, where μ is the reduced mass and σ the Lennard–Jones "diameter" of the pair. The points, with the exception of those for hydrogen and helium, indicate a straight line through the origin. Since the factor $1/\sqrt{\mu}$ is proportional to the average relative velocity and σ is a measure of the distance of closest approach, it

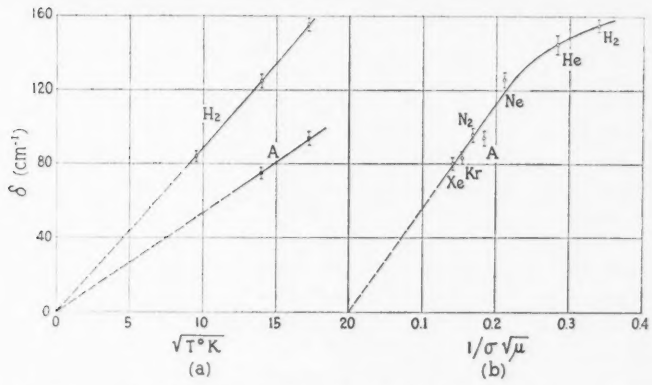


FIG. 7. Variation of the dispersion half width, δ, with (*a*) the temperature, T, and (*b*) the reduced mass, μ, and the Lennard–Jones "diameter", σ, of the collision pair.

might be argued that $1/\sigma\sqrt{\mu}$ is inversely proportional to the time spent in the region of interaction. The linear relationship in Fig. 7 (*b*) indicates that δ is inversely proportional to the duration of the interaction, and can be construed as a consequence of the Heisenberg uncertainty principle. Since the absorption profile shows only a small change over the range of densities used, the density dependence of δ is slight. Some measure of the density effect on the half width was obtained from the data for pure hydrogen at 85° K, where the rate of increase of δ was found to be about 4 cm^{-1} per 100 Amagat units of density. This is of the order of magnitude of ordinary collision broadening in infrared and microwave spectra of nonpolar molecules, and adds only a small amount to the large half width which is inherent in the process of induced absorption.

DOUBLE ROTATIONAL TRANSITIONS IN THE SPECTRUM OF HYDROGEN

At 85° K the $S(0)$ and $S(1)$ lines in normal hydrogen were observed at densities up to 118 Amagat, as shown in Fig. 1. At higher densities, where the maxima of these lines gave more or less complete absorption, two much weaker lines were observed on the high-frequency wing of $S(1)$ (Fig. 8). The maxima of these weak lines lie close to 941 and 1174 cm^{-1}, the calculated positions for the double transitions, $S(0)+S(1)$ and $S(1)+S(1)$, in which the two molecules in a binary collision perform rotational transitions simultaneously. Since the lines are superimposed on the high-frequency wing of the intense $S(1)$ line, their intensities were difficult to measure. In particular it was confirmed that the dispersion line shape for $S(1)$ no longer holds in this region as was noted above (Fig. 4); the intensity falls off more rapidly than the dispersion line form. Because of the difficulty in estimating the background intensity the binary absorption coefficients for the double transitions, given in Table II, cannot claim great accuracy; they should perhaps be considered as minimum values.

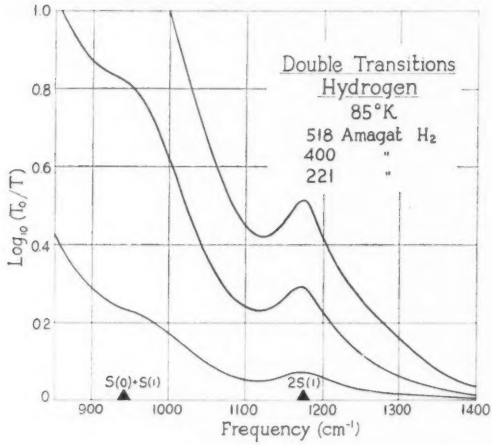


FIG. 8. The absorption of pure hydrogen at high gas densities in the region 900–1400 cm^{-1}, showing double rotational transitions.

TABLE II

Integrated absorption coefficients for double rotational transitions in hydrogen at 85° K

	$\tilde{\alpha}_1[S(1)+S(1)]$, $sec^{-1}\,cm^6$	$\tilde{\alpha}_1[S(1)+S(0)]$, $sec^{-1}\,cm^6$	$\tilde{\alpha}_1[S(0)+S(0)]$, $sec^{-1}\,cm^6$
Experimental	3.2×10^{-37}	0.7×10^{-37}	—
Theoretical	4.5×10^{-37}	2.2×10^{-37}	1.3×10^{-37}

The occurrence of the double rotational transitions depends on the anisotropy of the polarizability of the hydrogen molecule. If the quadrupole moment of molecule 1 induces a dipole moment in molecule 2, the dipole moment is modulated also by the rotational motion of molecule 2 if the polarizability is not completely isotropic. The calculation of this effect has been given by Van Kranendonk and Kiss (1959), using Volkmann's (1935) value of the anisotropy of the polarizability obtained from the depolarization of Rayleigh scattering in hydrogen. The calculated values of $\tilde{\alpha}_1$ are somewhat higher than those measured. The discrepancy must be partially due to the inaccuracy in the experimental measurement and partially to the error ($\pm 10\%$) in the value of the anisotropy of the polarizability.

THE DENSITY DEPENDENCE OF THE INTEGRATED ABSORPTION COEFFICIENT IN H_2–Xe MIXTURES

It was noted in Part I that the representation of the integrated absorption coefficient for H_2–Xe mixtures at room temperature as a function of density was of the form,

$$\int \alpha d\nu = \alpha_1 \rho_{H_2} \rho_{Xe} + \alpha_2 \rho_{H_2} \rho_{Xe}^2,$$

where $\alpha_1 = 1.74\times10^{-2}$ cm^{-2} Amagat^{-2} and $\alpha_2 = -5.5\times10^{-5}$ cm^{-2} Amagat^{-3}. The small cubic term indicates some effect of ternary collisions of the type Xe–H_2–Xe. Since the cubic term was not required in the case of other mixtures in the same density range, it was proposed that its presence for the H_2–Xe mixture might be due to molecular clustering in the neighborhood of the critical point (16° C) of xenon. The molecular clustering would lead to an abnormally high occurrence of Xe–H_2–Xe collisions in which some cancellation of the induced dipole would occur, leading to a negative ternary coefficient, α_2.

Experiments on the intensity of the absorption for H_2–Xe mixtures were therefore carried out at a series of temperatures 0°, 25°, 75°, and 98° C. The values obtained for the integrated binary and ternary coefficients are given in Table III.

TABLE III

Integrated absorption coefficients for H_2–Xe mixtures

T, °C	α_1, cm^{-2} Amagat^{-2}	α_2, cm^{-2} Amagat^{-3}
0	1.72×10^{-2}	-5.2×10^{-5}
25	1.74×10^{-2}	-5.5×10^{-5}
75	1.75×10^{-2}	—
98	1.70×10^{-2}	—

The values of α_1 are the same within the experimental error for all the temperatures. However, the ternary coefficient, α_2, has definitely a negative value at 0° and 25°, and is too small to be measured at 75° and 98° C. The evidence seems therefore to support the suggestion that the relatively large cubic term, $\alpha_2\rho_{H_2}\rho_{Xe}^2$, is due to molecular clustering in the neighborhood of the critical point. It might be noted that the isothermals of xenon over the pressure range used are nearly linear at 75° and 100° C, but depart very markedly from linearity at 0° and 25° C.

REFERENCES

CHISHOLM, D. A., MACDONALD, J. C. F., CRAWFORD, M. F., and WELSH, H. L. 1952. Phys. Rev. **88**, 957.
KISS, Z. J., GUSH, H. P., and WELSH, H. L. 1959. Can. J. Phys. **37**, 362.
KISS, Z. J. and WELSH, H. L. 1959. Phys. Rev. Letters, **2**, 166.
VAN KRANENDONK, J. and KISS, Z. J. 1959. Can. J. Phys. **37**, 1187.
VOLKMANN, H. 1935. Ann. Physik, **24**, 451.

THE MAGNETRON GAUGE: A COLD-CATHODE VACUUM GAUGE[1]

P. A. REDHEAD

ABSTRACT

A cold-cathode ionization gauge with axial magnetic field is described which is capable of measuring pressure in the range 10^{-3} to 10^{-12} mm Hg, and is primarily useful in the range above 5×10^{-10} mm Hg. The gauge is operated with an anode voltage of 6 kilovolts and a magnetic field of 1000 gauss. The relationship between ion current and pressure is linear in the pressure range 10^{-4} to 5×10^{-10} mm Hg; at lower pressures the relationship $i_+ = cP^{1.7}$ obtains. In the linear region the ion current per unit pressure is given by i_+ (amperes) $\simeq 9P$ (mm Hg) for nitrogen (i.e., a sensitivity about 45 times greater than obtainable with a thermionic triode gauge).

1. INTRODUCTION

It has been shown (Redhead 1958; Hobson and Redhead 1958) that a Townsend discharge in crossed electric and magnetic fields (Penning discharge) can be maintained at pressures as low as 10^{-13} mm Hg by suitable electron trapping in the discharge space. Moreover, it was shown that a Penning discharge gauge could be used to measure pressures below those measurable with a Bayard–Alpert gauge (Bayard and Alpert 1950). The Penning discharge is not limited by the X-ray effect which prevents measurements below about 5×10^{-11} mm Hg with a Bayard–Alpert gauge. The vacuum gauge described in these earlier papers was called an "inverted-magnetron gauge" since the anode was located on the axis and surrounded by a cylindrical cathode. The inverted-magnetron gauge is useful in the pressure range 10^{-3} to 10^{-13} mm Hg; its ion current – pressure relationship is non-linear and given by

$$i_+ = cP^n,$$

where i_+ is the positive ion current, c is a constant, and the exponent n lies between 1.10 and 1.15 at the normal operating point (anode voltage 6 kilovolts, magnetic field 2000 gauss).

The gauge described herein is based on similar design principles and has the geometry of a normal magnetron, i.e., the cathode is located on the axis and surrounded by a cylindrical anode. The magnetron gauge has certain advantages over the inverted-magnetron gauge for some applications, namely:

(*a*) the relationship between ion current and pressure, at the normal operating point (anode voltage 5–6 kilovolts, magnetic field 1000 gauss), is linear in the pressure range 10^{-3} to 5×10^{-10} mm Hg;

(*b*) the sensitivity of the magnetron gauge is higher than that of the inverted-magnetron gauge;

[1]Manuscript received June 18, 1959.
Contribution from the Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5383.

(*c*) a lower magnetic field is required, thus reducing the magnet size. The inverted-magnetron gauge is more suitable for measurements below 5×10^{-10} mm Hg.

The magnetron gauge has a sensitivity about 45 times greater than that of a standard Bayard–Alpert gauge (sensitivity 25 $(\text{mm Hg})^{-1}$, electron current 8 milliamperes). Since it contains no hot filament it is capable of measuring pressures of hydrogen and other active gases which would be decomposed by the hot filament of a thermionic ionization gauge.

2. CONSTRUCTION AND PROCESSING OF THE GAUGE

Figure 1 shows schematically the arrangement of the electrodes in the magnetron gauge. The cathode is shaped like a spool consisting of a short

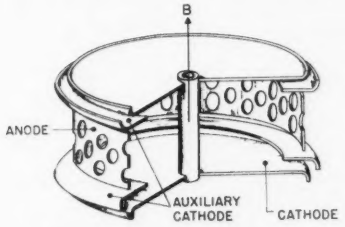


FIG. 1. Electrode arrangement of magnetron gauge.

cylinder to which are welded two circular end disks. The anode is a short, open-ended cylinder which is perforated to improve gas flow through the gauge. Two annular electrodes, the auxiliary cathodes, are inserted between the anode ring and the cathode end disks, and are maintained at cathode potential to shield the cathode end disks from high electric fields. The edges of the auxiliary cathode are rounded and electropolished to reduce field emission to a minimum. The gauge is enclosed in a cylindrical glass envelope with a flattened portion to allow insertion between the poles of a magnet. A conducting film is deposited on the inside surface of the glass envelope and is electrically connected to the auxiliary cathodes by a spring contact. This film acts as an electrostatic shield for the cathode end disks. Table I gives the salient electrode dimensions of a typical gauge.

TABLE I

Dimensions of magnetron gauge No. 25

Electrode	Symbol	Dimensions, mm
Anode radius	r_a	15
Anode length	l	14
Cathode radius	r_k	1.45
Cathode length		20

The electrodes are degreased, chemically cleaned, and then fired at 1050° C in vacuum for 1 hour. The electrodes are assembled and sealed into the

vacuum system rapidly after vacuum firing. The system is pumped and baked at 500° C for at least 18 hours. The gauges receive no further outgassing other than that caused by ion bombardment when the discharge is first established. After the above treatment there is no observable evolution of gas from the gauge at the lowest pressures so far obtained. The operation and degassing of the gauge is not affected if the system is let down briefly to air and the baking process subsequently repeated.

3. PERFORMANCE OF THE GAUGE

(*a*) *Ancillary Equipment*

Figure 2 shows the gauge and its associated circuits. The gauge is operated in an axial magnetic field of about 1000 gauss which is provided by a small

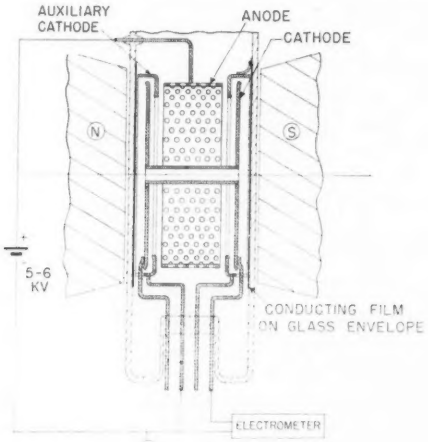


FIG. 2. Schematic diagram of magnetron gauge and associated circuits.

C-shaped permanent magnet. For the measurement of pressures greater than 10^{-9} mm Hg the anode voltage (5 to 6 kilovolts) may be derived from a simple unregulated power supply. For measurements below 10^{-9} mm Hg the anode voltage supply must be well regulated to prevent fluctuations in the supply voltage coupling to the electrometer through the interelectrode capacitance. A stabilized supply with 0.01% fluctuation for 20% change in line voltage has been found satisfactory (J. Fluke Co. Model 400BDB). At these pressures the gauge must be shielded from ambient light to prevent photocurrents to the ion collector. The positive ion current is measured by an electrometer in the cathode lead. The auxiliary cathodes are directly connected to ground; thus any field emission from these electrodes is not measured by the electrometer.

(*b*) *Calibration Method*

Measurements in the pressure range 10^{-3} to 10^{-10} mm Hg were made on continuously pumped systems. The lowest pressures were obtained by immers-

ing part of the system in liquid helium (Hobson and Redhead 1958). All measurements were made in systems constructed of Pyrex glass (Corning Type 7740). All pressures quoted in this paper are equivalent nitrogen pressures.

The magnetron gauge was calibrated directly against a Bayard–Alpert gauge (Bayard and Alpert 1950) in the pressure range where this gauge is operable (10^{-3} to about 10^{-10} mm Hg). For pressures less than 10^{-10} mm Hg an indirect method of calibration was employed.

Figure 3 shows schematically the apparatus used for the calibration of the magnetron gauge (M.G.) at pressures below those directly measurable with the

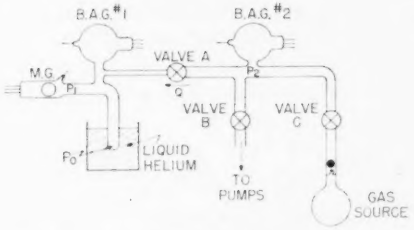


FIG. 3. Schematic diagram of system for calibration at very low pressures.

Bayard–Alpert gauge (B.A.G.). This system consisted of two main portions: volume 1, which contained the magnetron gauge, B.A.G.-1, and a finger which could be immersed in liquid helium; volume 2, which contained B.A.G.-2 and was isolated by three ultrahigh-vacuum valves (Bills and Allen 1955). Valve A controlled the conductance between volumes 1 and 2, valve B isolated the system from the pumps, and valve C was used to admit gas from a reservoir.

Once this system had been pumped to a pressure less than 10^{-10} mm Hg, valves A and B were closed. The finger was immersed in liquid helium with B.A.G.-1 turned off and the M.G. operating. Valve C was then opened and gas admitted to volume 2 until a steady pressure (P_2) of about 10^{-8} mm Hg was obtained. Valve A was then opened slightly and a steady flow (Q) of gas permitted to enter volume 1. Then,

$$Q = G_A (P_2 - P_1) \doteq G_A P_2 \qquad \text{(since } P_2 \gg P_1\text{)},$$

where G_A is the conductance of valve A. The steady value of P_1 obtained was determined by the input rate of gas and the sum of the pumping speeds of the magnetron gauge and the cold finger, thus

$$P_1 = \frac{Q}{S_{MG} + S_F},$$

where S_{MG} is the pumping speed of the magnetron gauge, and S_F the effective pumping speed of the cold finger. Now,

$$\frac{1}{S_F} = \frac{1}{S_T} + \frac{1}{G_T},$$

where S_T is the pumping speed of the cold surface at 4.2° K, and G_T is the conductance of the tubing between the cold surface and the magnetron gauge. It has been shown (Hobson 1959) that S_T is large compared with G_T in these measurements, then,

$$S_F \doteq G_T,$$

and since $G_T \gg S_{MG}$, then

$$P_1 = \frac{Q}{G_T} = \frac{G_A P_2}{G_T}.$$

With valve A set to a fixed conductance, P_1 is proportional to P_2. Various values of P_1 were obtained by adjusting valve C and calibration achieved by plotting the magnetron gauge current against P_2 as measured by B.A.G.-2. The value of G_A/G_T was obtained when P_1 was sufficiently high to be measured directly by B.A.G.-1.

(*c*) *Measured Characteristics*

The optimum operating point of the magnetron gauge is with an anode voltage (V_A) of 5 to 6 kilovolts and a magnetic field (B) of 1000 gauss. This point was chosen because the ion current – pressure relation is linear and the operation is most stable under these conditions. Figure 4 shows the relationship between ion current and pressure in the range 10^{-4} to 10^{-9} mm Hg, measured in a continuously pumped system by direct calibration against a B.A.G. These

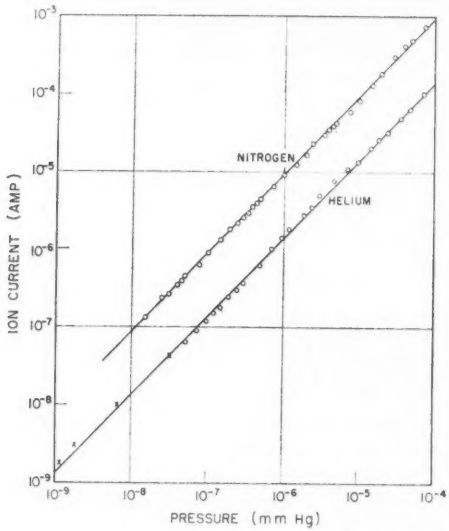


FIG. 4. Current–pressure relationship for M.G. No. 26 (with $V_A = 6$ kilovolts, and $B = 1070$ gauss) in the pressure range 10^{-9} to 10^{-4} mm Hg.

data were taken at $V_A = 6$ kilovolts, $B = 1000$ gauss. The measured points follow a straight line of unity slope within $\pm 10\%$ for nitrogen, and $\pm 20\%$ for helium. For nitrogen,

$$i_{+\ (\text{amperes})} \simeq 9P$$

where P is the pressure (mm Hg). The data for helium were obtained by taking the sensitivity ratio of the B.A.G. for helium and nitrogen as 0.209 (Schulz 1957). The sensitivity ratio of the M.G. for the same gases is given by

$$K_{He}/K_{N_2} = 0.15.$$

Operation of the gauge at other values of field and voltage gives non-linear pressure–current curves. For a linear relation between pressure and current V_A must be between 5 and 6 kilovolts, and $B = 1000$ gauss $\pm$ 10%.

The operating point must be carefully chosen to avoid instabilities in the discharge. Operation away from the optimum operating point results in low-frequency oscillations over limited pressure ranges and in sudden jumps in the pressure–current relation. These unstable regions are closely related to the striking characteristics which will be discussed in the next section.

The current–pressure relationship below 10^{-10} mm Hg is no longer linear, as is shown in Fig. 5 for helium. For pressures above about 5×10^{-10} mm Hg the

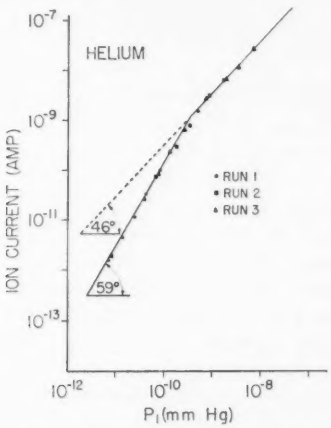


FIG. 5. Current–pressure relationship for M.G. No. 25 (with $V_A = 5$ kilovolts, and $B = 1060$ gauss) in the pressure range 10^{-8} to 10^{-11} mm Hg.

relationship is linear; at 5×10^{-10} mm Hg a sharp break occurs and the current–pressure relation below this pressure follows a power law given by

$$i_+ = cP^{1.7}.$$

The data of Fig. 5 were obtained by the indirect method of calibration described in Section 3(*b*).

The break in the characteristic of the magnetron gauge has been observed for various voltages and magnetic fields and different electrode dimensions, and appears to be intrinsic to this type of gauge geometry since the break is not observed in the inverted-magnetron gauge (Hobson and Redhead 1958).

High-frequency oscillation can be observed in the discharge at almost all pressures, the amplitude of these oscillations decreases with pressure, and they are almost immeasurably small at pressures below 10^{-10} mm Hg. The oscillations do not appear to affect the linearity of the ion current – pressure relation. At higher pressures care must be taken to ensure that the oscillations are not rectified by the electrometer to produce spurious readings.

The striking time for the gauge increases as the pressure is decreased; at a pressure of 10^{-9} mm Hg the striking time is about 10 seconds when the gauge is operated at the normal operating point ($V_A = 6$ kilovolts, $B = 1000$ gauss) and when the gauge is completely shielded from ambient light. At the lowest pressure obtained so far (about 10^{-13} mm Hg) the striking time may exceed 10 minutes when the gauge is completely shielded. It is usually necessary to irradiate the gauge with ultraviolet light from a mercury lamp to ensure rapid starting at these pressures; the radiation is removed as soon as the gauge has struck.

The pumping speed of the gauge is approximately 0.15 liter/second for helium and 1 liter/second for air at the normal operating point. One gauge constructed entirely of titanium showed a pumping speed to air of about 8 liters/second.

4. THE BEHAVIOR OF THE DISCHARGE

(*a*) *Electron and Ion Trajectories*

The electron and ion trajectories, in the absence of space charge, will be described briefly. Electrons cannot escape from the discharge in an axial direction because of the end plates at cathode potential. Provided the magnetic field exceeds the cutoff field, which is the only case of interest in crossed-field discharges, the electrons are trapped radially by the magnetic field. Electrons will follow a hypocycloid path around the cathode while simultaneously oscillating back and forth in the axial direction. Electrons liberated from the cylindrical portion of the cathode surface perform one arc of a cycloid and return to the cathode unless they suffer a collision with a gas molecule during this first arc. The probability of an electron undergoing a collision in the first cycloidal arc is termed "the cathode escape probability" (C_k). Electrons suffering inelastic collisions are assumed to return to thermal velocities and to start their cycloidal motion once more from the point of collision. Electrons which have escaped from the cathode or are liberated within the interelectrode space will circle the cathode in a hypocycloid generated by a circle of diameter D rolling on a circle of constant radius. The electron is trapped indefinitely in this hypocycloidal orbit until it eventually suffers a collision with a gas molecule. At each inelastic collision the electron returns to near zero energy and restarts its cycloidal motion from a point closer to the anode. Thus

electrons can only progress towards the anode by successive collisions with gas molecules.

The maximum energy of an electron in its hypocycloidal orbit is given by

$$V_{max} = ED \tag{1}$$

where E, the electric field $= V_A/\{r \log(r_a/r_k)\}$,
V_A is the anode voltage,
r_a the anode radius,
and r_k the cathode radius.

When V_{max} is less than the ionization potential of the gas (V_i) the electron is no longer capable of causing ionization. When $V_{max} = V_i$, $r = r_1$. Thus for $r > r_1$ no ionizing collisions are possible in the absence of space charge.

The magnetic field in the magnetron gauge is considerably below the cutoff value for positive ions, thus the ions formed by inelastic electron collisions move directly to the cathode. The transit time for positive ions is considerably shorter than for electrons, thus the space charge in the magnetron gauge is electronic except at very high pressures ($P > 10^{-2}$ mm Hg).

(*b*) *Striking Characteristics*

An approximate theoretical treatment of the striking characteristic of crossed-field discharges has been presented previously (Redhead 1958), in which it is assumed that the electric field is constant over any one cycloidal arc of the electron trajectory. The following discussion is based on this previously presented theory.

The conditions for the start of the discharge are governed by the Townsend breakdown criterion which, for non-uniform fields, can be written:

$$\int \alpha dr = \log\left(1+\frac{1}{\Gamma}\right) \tag{2}$$

where α, the first Townsend coefficient, is the number of ions produced by an electron per unit distance advance in the direction of the electric field, and Γ, the effective third Townsend coefficient, is the number of electrons re-entering the avalanche process per ion leaving the discharge (i.e., $\Gamma = C_k\gamma$, where γ is the secondary electron emission coefficient of the cathode for positive ions).

Figure 6 shows the measured striking characteristics of a magnetron gauge for nitrogen in the pressure range 10^{-9} to 10^{-5} mm Hg. Figure 7 shows the corresponding characteristics for helium.

The V_A-B diagrams may be divided into three regions by the lines $r_1 = r_a$ and $r_1 = r_k$ (see Figs. 6 and 7). r_1, the radius at which the maximum energy of the electron in its hypocycloidal orbit is equal to the ionization potential of the gas, is given by

$$r_1^2 = \frac{2}{\eta V_i}\left(\frac{E_k r_k}{B}\right)^2, \tag{3}$$

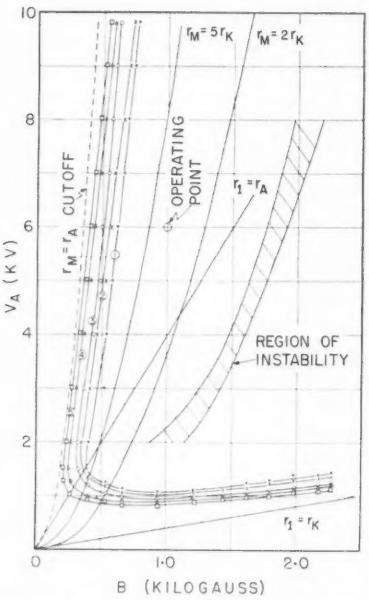


FIG. 6. Striking characteristics of M.G. No. 25 for N_2.
Curve 1, $P = 2.5\times10^{-9}$ mm Hg; curve 2, $P = 2.5\times10^{-8}$ mm Hg; curve 3, $P = 3.0\times10^{-7}$ mm Hg; curve 4, $P = 2.5\times10^{-6}$ mm Hg; curve 5, $P = 2.5\times10^{-5}$ mm Hg.

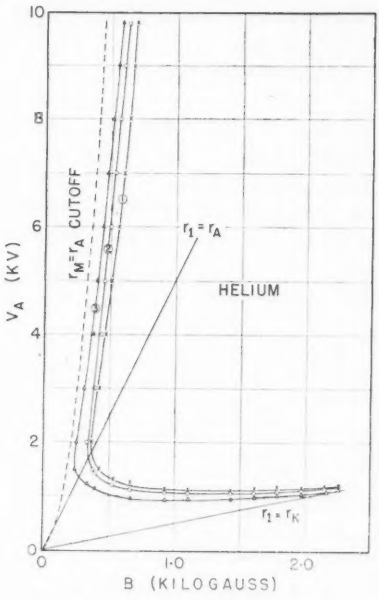


FIG. 7. Striking characteristics of M.G. No. 25 for He.
Curve 1, $P = 1.9\times10^{-7}$ mm Hg; curve 2, $P = 1.9\times10^{-6}$ mm Hg; curve 3, $P = 1.9\times10^{-5}$ mm Hg.

where η is the charge-to-mass ratio of the electron,
E_k the electric field at the cathode surface,
and B the magnetic field.

When $r > r_1$, electrons do not have sufficient energy to cause ionizing collisions. In region (a), above the line $r_1 = r_a$ (see Fig. 6), electrons are capable of entering into ionizing collisions with gas molecules throughout the entire interelectrode space. In region (b), between the lines $r_1 = r_a$ and $r_1 = r_k$, ionizing collisions are only possible within the volume bounded by the cathode surface and the surface $r = r_1$ ($r_k < r_1 < r_a$). In region (c), below the line $r_1 = r_k$, ionizing collisions cannot occur anywhere in the interelectrode space.

The above analysis ignores the effects of space charge and the distortion of the electric field caused by the cathode end plates. It can be shown that the striking characteristics are only slightly altered by these two effects. Space charge in the prebreakdown conditions is extremely small and does not build up to any extent until after the Townsend breakdown criterion is satisfied. The end effects have also been shown to be small by visually observing the start of the discharge in a darkened room at pressures of about 10^{-8} mm Hg. The discharge is observed to strike in a thin lamina, with well-defined boundaries,

at the middle of the cathode cylinder where the electric field is greatest and where the field distortion by the end plates is least. The space-charge-free potential distribution in these gauges has been found from measurements with an analogue computer which show that the field in a plane through the middle of the cathode differs insignificantly from the theoretical field between coaxial cylinders of infinite length.

It can be seen from Figs. 6 and 7 that the line $r_1 = r_a$ passes through the knee of the curves and divides the measured striking characteristics into two branches. The upper branch of the characteristic lies entirely in region (*a*), where ionizing collisions can occur anywhere in the interelectrode space. This branch is very closely parabolic, in qualitative agreement with the theory of breakdown in crossed fields presented previously (Redhead 1958), which gives for the striking voltage on the upper branch,

$$V_s = \frac{k_1 B_s^2}{\log[1+(1/\Gamma)]}, \tag{4}$$

where

$$k_1 = \frac{3\eta}{8}(r_a^2 - r_k^2)\log(r_a/r_k),$$

B_s is the striking magnetic field,

and Γ, the effective third Townsend coefficient $= f(V_A, B, P)$.

When the dominant secondary processes occur at the cylindrical portion of the cathode surface, Γ is given by

$$\Gamma = C_k \gamma, \tag{5}$$

where γ is the secondary electron emission coefficient of the cathode surface for positive ions, and the cathode escape probability (C_k) may be approximated by (see Redhead 1958)

$$C_k = \frac{8E_k}{B^2\lambda} = k_2 \frac{V_A}{B^2} P, \tag{6}$$

where λ is the electron mean free path,

and k_2 is a constant.

Since V_A/B^2 is nearly constant for the upper branch of the striking characteristic any changes in Γ will be caused by variations of γ with ion energy; this effect is quite small.

The above analysis assumes that the electric field is constant over any one cycloidal arc of the electron trajectory. The maximum radius (r_m) that an electron, released from the cathode surface, can reach in its first swing outward from the cathode has been calculated, and lines of constant r_m/r_k plotted in Fig. 6. It can be seen that the upper branches of the striking characteristic lie in the region where $r_m > 5r_k$ (or $r_m > \frac{1}{2}r_a$). Thus the assumption of a constant electric field over the cycloidal arc is a very poor approximation for the upper branch. This analysis correctly predicts the general shape of the upper branch

and its movement with pressure, but it is not capable of giving quantitative information on the secondary processes involved.

The lower branch of the striking characteristics falls in region (*b*) of the $V_A - B$ diagram; on this branch ionization is only possible in a limited portion of the interelectrode space surrounding the cathode. The theory gives, for the lower branch,

$$\log\left(1+\frac{1}{\Gamma}\right) = \frac{3E_k r_k}{2V_i}\log\frac{3/2}{1+(\eta V_i/4)(B/E_k)^2}. \tag{7}$$

A numerical solution of this equation, with Γ assumed constant, shows that for $B > 10^3$ gauss, the characteristic rises slowly with B. However, on the lower branch Γ will change considerably since V_A is approximately constant, and B varies over a wide range. Thus Γ decreases with increasing B, which tends to maintain the striking voltage (V_s) constant. At higher pressures, and thus larger Γ, the numerical analysis indicates that V_s will increase more rapidly with B, as is found experimentally (see Figs. 6 and 7).

(*c*) *Running Characteristics*

When the magnetron gauge is operated at constant anode voltage and the magnetic field increased from zero, the ion current is observed to increase rapidly, once the striking magnetic field is exceeded, and then become approximately constant. As a critical magnetic field is reached the ion current decreases instantaneously by a factor of about five and then remains constant at this new level as B is increased. When the field is reduced, an upward step in ion current is observed at a field slightly lower than when increasing B. The region where this step occurs is indicated in Fig. 6 as the region of instability. The points at which the steps occur lie on a parabola just below the boundary between regions (*a*) and (*b*) and are only slightly affected by pressure.

Once the discharge has struck, the current is limited by space charge. The effect of the space charge, which is electronic at all pressures below about 10^{-2} mm Hg, is to reduce the field at the cathode surface and thus decrease the probability of escape of the secondary electrons. The current through the discharge adjusts itself until the field at the cathode surface is such that the effective secondary emission coefficient ($\Gamma = \gamma \times$cathode escape probability) satisfies the Townsend breakdown criterion. As the magnetic field is increased and the operating point moves into region (*b*) of the $V_A - B$ diagram the space charge increases the field at the anode above the zero space-charge value so that electrons are still capable of ionizing near the anode surface. At some value of B the electric field at the cathode becomes so small that electrons leaving the cylindrical part of the cathode never acquire sufficient energy to ionize before striking the cathode once more. For B greater than this critical value the discharge is maintained only by secondary processes from the cathode end plates. Since most of the ion current strikes the cylindrical part of the cathode, at this critical magnetic field there is a sharp reduction in the number of secondaries re-entering the discharge and a corresponding drop in the ion current.

The magnetic field has little effect on the probability of a secondary electron leaving the end plates and entering the discharge since most of these electrons are moving parallel to the magnetic field. Thus Γ becomes independent of B once the critical field (B_c) has been exceeded. This explains the constancy of ion current with increasing B, when $B > B_c$. The hysteresis in B_c is caused by the different space-charge conditions existing on either side of the region of instability.

5. SUMMARY

The magnetron gauge is capable of measuring pressure from 10^{-3} to 10^{-12} mm Hg. It has the following advantages over a thermionic triode gauge:

(*a*) It is not limited by the X-ray effect and thus can measure lower pressures.

(*b*) There is no hot filament to cause decomposition of hydrogen or oxygen, or to evolve carbon monoxide.

(*c*) The sensitivity of the magnetron gauge in the linear region is about 45 times greater than that of a standard Bayard–Alpert gauge.

The pumping speed of the magnetron gauge is greater than that of the triode gauge by approximately the same ratio as the ion currents. Since the range of operating voltage and field is small, the pumping speed cannot be readily changed as in the triode gauge, thus the magnetron gauge cannot be used in experiments where the pumping speed must be kept small.

The magnetron gauge is most useful in the pressure range 10^{-4} to 5×10^{-10} mm Hg where its sensitivity is high and its calibration linear. At a pressure of 5×10^{-10} mm Hg the pressure–current relation becomes non-linear and if the gauge is used below this pressure it must be calibrated for the gas to be used by the method described in Section 3(*b*). For measurement of these low pressures the inverted-magnetron gauge (Hobson and Redhead 1958) is preferable since there is no break in its calibration curve.

The measured striking behavior of the magnetron gauge is in qualitative agreement with the simple ballistic theory presented previously.

ACKNOWLEDGMENTS

The author is indebted to J. P. Hobson and E. V. Kornelsen for the measurements at liquid helium temperatures, and to R. D. Cottee and A. W. Pye for constructing the gauges and vacuum systems.

REFERENCES

Bayard, R. T. and Alpert, D. 1950. Rev. Sci. Instr. **21**, 571.
Bills, D. G. and Allen, F. G. 1955. Rev. Sci. Instr. **26**, 655.
Hobson, J. P. 1959. Can. J. Phys. **37**, 300.
Hobson, J. P. and Redhead, P. A. 1958. Can. J. Phys. **36**, 271.
Redhead, P. A. 1958. Can. J. Phys. **36**, 255.
Schulz, G. J. 1957. J. Appl. Phys. **28**, 1149.

FRONTAL PERTURBATION OF A TROPOSPHERIC SCATTER PATH[1]

D. R. Hay[2] and G. E. Poaps

ABSTRACT

A 2-year study of 500 Mc/s radio transmissions over an 85-mile path near Ottawa has shown that the signal fading rate rises well above the diurnal maximum when the transmission path is perturbed by a weather front. The fading rate reaches a maximum when the upper boundary of the frontal zone is 3000 feet above the center of the radio path, and the fading rate remains high as long as any part of the frontal zone is between the surface and 3000 feet. A similar relationship between frontal position and signal fading rate is found for Arctic, Maritime, or Polar fronts in the Ottawa area, but the Arctic fronts provide the clearest definition. There is some indication that high fading rate also may be associated with horizontal layers of contrasting humidity at low levels in the troposphere.

I. INTRODUCTION

A well-known characteristic of radio transmission at ultrahigh frequencies (300 to 3000 Mc/s) is its dependence upon the weather along the radio path. When the radio transmitter and receiver are located either within or just beyond the line of sight, the signal transmission loss may be estimated with reasonable certainty if the refractivity profile of the local air mass is known. But the computation of transmission loss becomes a complicated problem when the receiver is well beyond the optical horizon of the transmitter. There appear to be two reasons for this complexity: First, the long radio path may pass through two or more different air masses, and each has its own transmission characteristics. The second reason is even more fundamental, in that the physical process by which the signal travels from the transmitter to the receiver is not clearly understood. At least three theories on the process of signal transmission have been suggested (see, for example, Gordon 1955; Carroll and Ring 1955; and Northover 1955). The principle that is common to these theories is that the radio signal from the transmitter is reflected or scattered within the troposphere and in the forward direction, towards the distant receiver.

The results of the experiment described in this paper suggest a new method for studying the physical process of tropospheric scatter transmission. They show that the radio signal path is disturbed strongly by a local weather front, whose presence causes the rate of fluctuation in the received signal to increase, and the mean amplitude of the signal to decrease. Further, the perturbation of the radio path persists as long as any part of the frontal zone is at a pre-

[1]Manuscript received July 30, 1959.
Contribution from the Defence Research Telecommunications Establishment, Defence Research Board, Ottawa, Canada. This work was carried out under Project D48-28-01-03.
[2]Now at the Department of Physics, University of Western Ontario, London, Ontario.

ferred height above ground. Thus it should be possible to examine the process of signal transmission if information on the physical structure of the perturbing front is known.

It is the purpose of this paper to describe the experiment that relates the frontal position to the perturbation of the U.H.F. radio path. No attempt has been made to examine the physical structure of the transmission path in detail. The information on the frontal positions has been drawn from synoptic weather charts and local cloud observations, and the experimental results for 1 year have been treated statistically to compensate for errors in interpolation of the frontal position.

II. RADIO SIGNAL FLUCTUATION

The experimental radio transmissions were carried out over an 85-mile path near Ottawa, Canada. A 500 Mc/s signal was transmitted from a 28-foot paraboloidal reflector antenna at Chalk River and received by a 20-foot antenna of similar construction near Ottawa. Both antennas were directed horizontally towards each other. The vertical profile along the transmission path, as illustrated in Fig. 1, shows that the radio horizon is within 28 miles

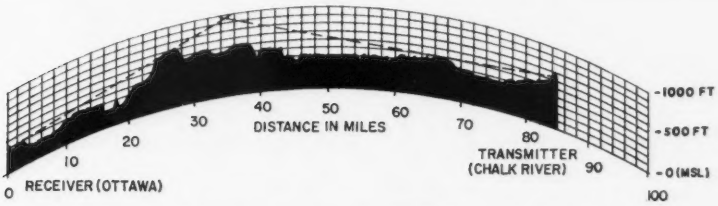


FIG. 1. Vertical profile along the Chalk River – Ottawa path.

of either antenna. The radius of curvature of the sea level in this diagram is four-thirds of the true radius, to simulate the refraction of a standard atmosphere. It will be seen that the horizon planes intersect about 400 feet above the surface near the center of the path.[3]

Typical scatter transmission occurred between Chalk River and Ottawa, in that the received signal amplitude fluctuated or faded continuously. These fluctuations are illustrated in the sample records in Figs. 2(*a*) and 2(*b*). The amplitude variation was irregular, and the range of variation exceeded 27 decibels. The rate of fading[4] in the signal amplitude generally changed in an orderly way as observed on equipment designed to record the fading rate directly and continuously (Day and Psutka 1956). Typical recordings obtained with this equipment are shown in Fig. 3. The normal variation in fading rate of the 500 Mc/s signal for a 2-day interval is illustrated in Fig. 3(*a*). Here,

[3]Further details on this experiment are given elsewhere in published reports (Hay and Poaps 1959) and (Hay, Day, and Maynard 1958).

[4]It is customary to define signal fading rate as the number of positive crossings of the mean signal level by the envelope of the signal, in a prescribed time interval. In this experiment, both the prescribed time interval and the interval over which the mean level is taken are 60 seconds.

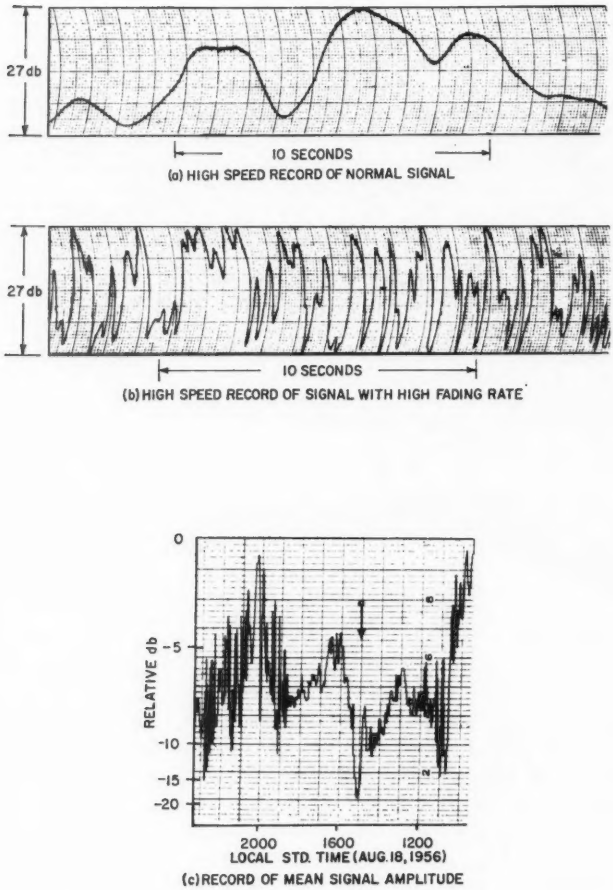


FIG. 2. Sample records of the time variation in the 500 Mc/s signal amplitude at Ottawa.

the maximum fading rate was about 20 per minute, during the early afternoon; the minimum fading rate occurred shortly after midnight, when less than five fluctuations per minute were indicated.[5]

Soon after the Chalk River – Ottawa transmissions began it became apparent that the normal pattern of change in fading rate was disturbed for

[5]The brief intervals of zero fading rate that are 12 hours apart are due to periodic shutdown of the Chalk River transmitter.

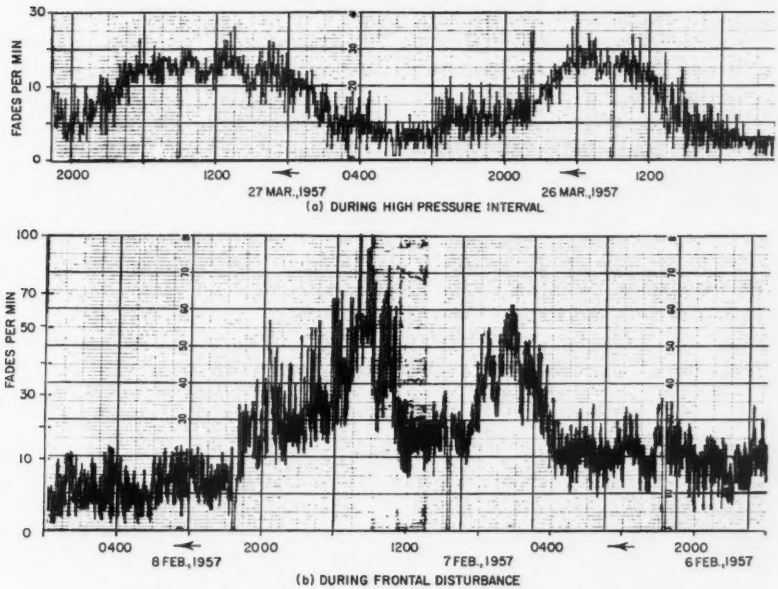


FIG. 3. Records of the fading rate in the 500 Mc/s signal at Ottawa.

several hours out of every 3 to 5 days. This disturbance of the normal signal transmission was shown in two ways. The first indication was an increase in signal fading rate, even above the maximum of the diurnal change in Fig. 3(*a*). The record of Fig. 2(*b*) represents the signal that was received during such an interval. A continuous plot of signal fading rate through a disturbed interval is illustrated in Fig. 3(*b*). Here it is seen that the fading rate was as high as 100 per minute, and that the fading rate remained high for 4 hours or longer.

The second indication of abnormality in signal transmission is illustrated in Fig. 2(*c*). This record represents the change in mean signal amplitude before, during, and after a period of disturbance. Note that the time scale of Fig. 2(*c*) is greatly compressed, and that the fluctuations shown in Figs. 2(*a*) and 2(*b*) have been eliminated in 2(*c*) by electrical integration over several minutes. The important feature of Fig. 2(*c*) is the marked depression in mean amplitude (at the arrow). Such depressions in received signal occurred whenever the signal fading rate was near a maximum during a disturbed interval.

III. DISTURBED TRANSMISSIONS

In the following analysis only those intervals in which signal transmissions were disturbed have been considered. Further, the criterion selected for disturbed transmission is the abnormally high fading rate, ***not*** the depression in

mean signal amplitude. Examination of the diurnal change in fading rate for undisturbed transmissions has shown that the maximum fading rate varied throughout the year. Table I illustrates this variation. Thus, for any interval during a given month in which the fading rate of the 500 Mc/s signal exceeds the indicated maximum, the transmission is assumed to be disturbed.

TABLE I

Maximum fading rate for undisturbed transmissions

Month	Maximum fading rate	Month	Maximum fading rate
January	20	July	15
February	19	August	15
March	18	September	16
April	18	October	18
May	17	November	20
June	16	December	21

The period from February 25, 1957, to February 25, 1958, has been selected for the analysis of signal fading rate. Although the number of useful recording days was reduced from 366 to 306 by equipment failures, these occurred irregularly, and hence the useful records are representative of the full year. One hundred and seventy-eight intervals of disturbed transmission were recorded, and the prominent weather features along the radio path at these times were examined as described below.

Some indication of the relationship between disturbed signal transmissions and specific features of the weather may be found in published accounts elsewhere. Prior to 1938, Hull found that the mean amplitude of a 60 Mc/s signal as received at a distance of 90 miles generally was depressed during periods of precipitation along the radio path (Friend 1945). The rate of fluctuation of the signal increased also at the same time. Pickard and Stetson (1947) reported at a later date that a prominent signal depression accompanied the appearance of a weather front in a long radio path. The depth of depression was greatest when the front made a large angle with the radio path. In Japan, it has been found that cold weather fronts in a 75-mile radio path accompanied a greater variation in signal amplitude than did warm fronts (Hirao, Uyesugi, and Tao 1952); the extent of the variation was greater at night than in the daytime. Experiments on radio paths up to 150 miles long in Germany have indicated that the depth of signal depression varied for the different types of weather fronts that disturbed the radio paths (Lauter and Klinker 1954; Klinker 1956). Randall (1954) in the United States has reported a signal amplitude enhancement prior to the depression, when a long radio path is perturbed by a cold weather front. Thus, an increase in signal fading rate and a depression in signal amplitude in V.H.F. and U.H.F. scatter transmissions have been associated with weather fronts in the radio path. However, little further information on this relationship is available.

IV. INITIAL ANALYSIS

Obvious irregularities in the local weather were examined first, for their relationship to disturbances in the radio transmission. All of the 178 periods of abnormally high fading rate have been included in this analysis. Information on the local weather was obtained from hourly observations at weather stations near the radio path.[6] That precipitation alone has little effect upon the signal transmission soon was apparent. Only 12% of the maxima in fading rate occurred when rain fell on a major part of the radio path, and only 6% of the maxima when snow fell.

The relationship between disturbed transmissions and air pressure minima at the radio path was little closer. It will be recalled that centers of low pressure generally are associated with strong vertical air currents. Figure 4(*a*)

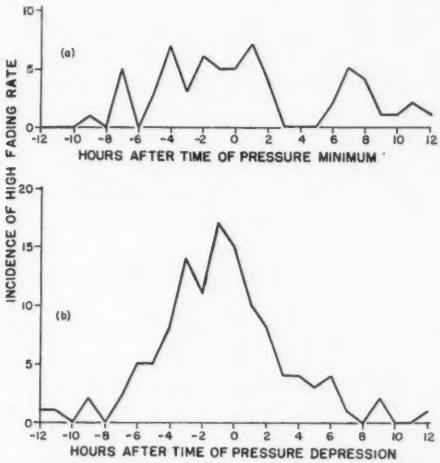


FIG. 4. Relation between disturbances in the 500 Mc/s signal transmission and the drop in air presssure at the radio path.

compares the time of the maximum in fading rate with the time of a minimum in surface pressure near the center of the radio path. In all, 102 well-defined minima in air pressure were recorded during the 306 days of signal transmission. But only 62 of these minima were accompanied within 12 hours by abnormally high fading rate in the radio signal.

A higher degree of correlation was found between shallow depressions in surface air pressure and disturbance of the signal transmission. These drops of some 2 millibars in air pressure with respect to the normal trend lasted

[6]Surface weather stations, operated by the Department of Transport, are located at Ottawa (Uplands), at Killaloe (32 miles southwest of the center of the radio path), and at North Bay (185 miles northwest of the Ottawa site). A radiosonde station is located at Maniwaki, Quebec (68 miles north of the receiver site).

from 1 to 2 hours. They were generally accompanied by a weather frontal surface. In 306 days, 120 shallow pressure depressions were recorded and the signal fading rate reached a maximum within 12 hours in all but two instances. This relationship is illustrated in Fig. 4 (*b*). The fact that the peak in this distribution is not at the zero time axis may be due to the separation between the radio path and the Killaloe weather station where the air pressure was recorded.

The initial analysis, then, suggests that it is primarily the structure of the frontal zone and not the associated precipitation or the vertical air movements at the center of low pressure that leads to a disturbance in signal transmission. This suggestion is borne out by a closer examination of the frontal heights above the radio path at the time of signal disturbances.

V. FRONTAL HEIGHT AND DISTURBED TRANSMISSIONS

The continuous changes in the height of weather fronts above the radio path have been deduced from published charts of frontal contour analysis.[7] Information on the orderly ascent or descent of clouds that lay just above the front (Grant 1944) has been obtained from hourly observations at the weather stations near the radio path to assist with the interpolation of frontal heights at low levels. By convention, the frontal height is measured from the upper (or warm air) boundary of the frontal zone. The full depth of the frontal zone is about 3000 feet, the depth depending upon the type and height of the front (Godson 1951; Sawyer 1955). Because the frontal zone is inclined at a shallow angle to the horizontal, the frontal height has been estimated for the center of the radio path.

When the frontal data are treated statistically for the whole period of the experiment, three weather fronts generally are assumed to exist above the radio path. However, at any time this number may decrease according to the local weather conditions. Polar fronts usually appeared near the surface in midsummer, Arctic fronts in midwinter, and Maritime fronts at other times of the year. Whenever two or three fronts occurred simultaneously overhead in the lower troposphere the mean vertical separation between the upper boundaries of the fronts was 6000 feet.[8] Thus, when the full interval of 306 days is considered statistically no more than one type of front is assumed to appear within 6000 feet of the surface.

The heights of weather fronts above the middle of the radio path during intervals of high signal fading rate are indicated in Fig. 5. All of the 178 intervals of disturbed transmission are represented in these graphs. Of these, 154 occurred while a weather front was within 6000 feet of the surface. The F.C.A. charts indicated either no front overhead or a front above 6000 feet for the other 24 intervals. It should be added here that the F.C.A. charts

[7]Frontal contour analysis (F.C.A.) charts are issued twice daily by the Central Analysis Office of the Meteorological Branch, in Dorval, Quebec. They are based upon the three-front model used by Canadian meteorologists (Anderson, Boville, and McClennan 1955; Penner 1955).

[8]The Polar front is uppermost, and the Arctic front is the lowest when the three appear together over the observer.

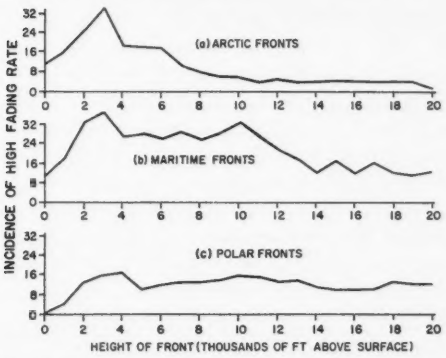

(a) ARCTIC FRONTS
(b) MARITIME FRONTS
(c) POLAR FRONTS
INCIDENCE OF HIGH FADING RATE
HEIGHT OF FRONT (THOUSANDS OF FT ABOVE SURFACE)

FIG. 5. Relation between disturbances in the 500 Mc/s signal transmission and the frontal height above the center of the radio path.

also indicated 24 weather fronts below 6000 feet during the period of the experiment when no disturbance in the signal transmission was recorded. Statistics for each of the three types of fronts found in the Ottawa area are plotted separately in Fig. 5. In Fig. 5(*c*), it will be seen that the incidence of disturbed transmissions is at a maximum when the Polar front is at 4000 feet. Figure 5(*b*) indicates two maxima in the incidence of disturbed transmissions when Maritime fronts alone are considered; one is for the frontal height of 3000 feet and the other for a height of 10,000 feet. In Fig. 5(*a*), the incidence of transmission disturbance reaches a well-defined maximum for Arctic fronts at 3000 feet, and the frequency decreases to a somewhat lower value until the Arctic fronts rise above 6000 feet. For frontal heights above this level, disturbed radio transmissions become relatively infrequent.

VI. REMARKS

An explanation of the results given in the previous section may be based upon the known physical structure of a weather front. Godson (1951) has examined the data obtained from many radiosonde ascents in Canada, and Sawyer (1955) has analyzed the results of flight soundings through fronts in the United Kingdom. They have concluded that there is no universal model of a front, but that certain features are common to most. For example, the average slope of a cold front is 1 in 50 and of a warm front is 1 in 250. The vertical thickness of the frontal zone may be between 1000 and 5000 feet at lower levels of the troposphere. Generally, the air in the frontal zone is very dry except at levels below 2000 or 3000 feet. Palmen and Newton (1948) have found that the frontal zone possesses a strong horizontal wind shear. A study of balloon ascents through a well-defined frontal zone has shown, also, that the vertical wind shear is discontinuous at the frontal zone boundaries (Reed 1957).

These features of a frontal cross section are consistent with the results of microwave refractometer soundings (Crain, Gerhardt, and Williams 1954).

The refractive index profiles show abrupt changes in the lapse rate of refractivity at the upper and lower boundaries of the frontal zone where the wind shear is discontinuous. And within the frontal zone where the air is relatively dry, the lapse rate in refractivity is approximately uniform but considerably different from the lapse rates in the air above and below the front.

It is suggested here, then, that the upper boundary of the frontal zone is the chief perturbing region as the front moves through the radio transmission path. The remainder of the frontal zone presents a perturbation of lesser magnitude.[9] The frontal zone alters the radio scattering process in the troposphere so that the signal fading rate is increased above the normal diurnal maximum and the transmission efficiency is reduced. In Fig. 5(*c*), the Polar front is the only one above the radio path when a front is within 6000 feet of the surface. When the Polar front is at higher levels, the Maritime front and possibly the Arctic front appear below it also in the radio path. Figure 5(*b*) indicates two preferred heights of the Maritime front for a disturbance in the signal transmission; and when this front has risen above 14,000 feet the incidence of disturbed transmissions decreases to an approximately constant value. Thus, it appears that the Maritime front is the cause of transmission disturbance as it rises through the first 4000 feet of the atmosphere, but the disturbance when the Maritime front is at 10,000 feet is due to the Arctic front below it (now within 4000 feet of the surface). Finally, in Fig. 5(*a*), it is suggested that the high fading rate results from perturbation of the radio path by the Arctic front alone.[10] Perturbation of the radio path becomes minor when the Arctic front is higher than 6000 feet.

Figure 6 illustrates the suggested explanation of radio path disturbance. It is consistent with the experimental results given above to assume that only

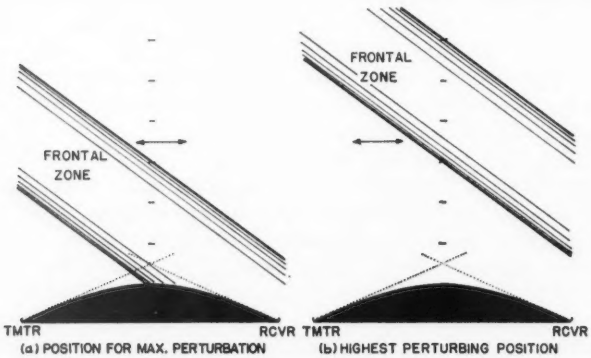


FIG. 6. Vertical section through the radio path and the perturbing front.

[9]It is reasonable to expect a more abrupt change in the lapse rate of refractivity at the upper boundary of the frontal zone than at the lower boundary, because the air is warm (and relatively moist) above this zone but cool (and drier) below.

[10]The Maritime front is now 6000 feet above the Arctic front, and the Polar front is 6000 feet above the Maritime front.

the lowest of the three fronts causes a significant perturbation.[11] The diagram represents a vertical section through the radio path when it is occupied by a weather front. The front, which extends perpendicularly to the plane of the diagram, is somewhat idealized here for the sake of illustration, and only the lowest front is considered. The vertical depth of the frontal zone is assumed to be 3000 feet.[12] Figure 6(*a*) represents the position of the front when the signal fading rate is at a maximum. Here, the upper boundary is about 3000 feet above the surface. Because the fading rate remains at a high level for frontal heights between the surface and 6000 feet, it is only for frontal positions above that shown in Fig. 6(*b*) that perturbation of the radio path becomes negligible. Note that the lower boundary in 6(*b*) is 3000 feet above the surface. The frontal motion is to the right at about 20 miles per hour if it is a cold front, or to the left at a lesser speed if it is a warm front. Thus, the lowest 3000 feet of the troposphere, and especially the 3000-foot level, appear to be critically important in radio transmissions over an 85-mile path.

This description of frontal perturbation has assumed a simplified picture of the troposphere. That other sources of perturbation may be present is suggested by the fact that the incidence of transmission disturbance does not vanish in Fig. 5(*a*) when all fronts are higher than 6000 feet. Twenty-four intervals of this type were recorded during the experiment. An examination of the radiosonde data from Maniwaki for these intervals showed a prominent irregularity in the lapse rate of temperature and relative humidity within 5000 feet of the surface in 13 instances. These irregularities generally represent horizontal layers of contrasting humidity in the air. Since horizontal stratification of this type often follows the curvature of the earth's surface for 100 or 200 miles (Bauer, Meyer, Prohaska, and Wilson 1956; Crain 1955), it is probable that such layers also contributed to the disturbance of signal transmissions.

ACKNOWLEDGMENTS

The authors are indebted to the Director of the Meteorological Branch, Department of Transport, and to the Officers-in-Charge of the weather stations at Maniwaki, Killaloe, North Bay, and Ottawa for supplying weather information between 1955 and 1958. They also are very grateful to Dr. W. Godson, Superintendent of the Atmospheric Research Section of the Meteorological Branch, and to Mr. J. Leaver, Officer-in-Charge of the Central Analysis Office at Dorval, for helpful advice and comments on the meteorology.

REFERENCES

Anderson, R., Boville, B. W., and McClennan, D. E. 1955. Quart. J. Roy. Meteorol. Soc. **81**, 588.

Bauer, J. R., Meyer, J. H., Prohaska, J. T., and Wilson, F. A. 1956. U.R.S.I. Meeting, University of California, Berkeley.

Carroll, T. J. and Ring, R. M. 1955. Proc. Inst. Radio Engrs. **43**, 1384.

[11]This assumption agrees also with the theory of radio signal scattering by atmospheric turbulence (Gordon 1955) since the scattering coefficient of turbulence is inversely proportional to the fourth power of the angle of scatter.

[12]Note that the vertical scale in this diagram is exaggerated about 100 times. The vertical markers represent 1000-foot intervals.

CRAIN, C. M. 1955. Proc. Inst. Radio Engrs. **43**, 1405.
CRAIN, C. M., GERHARDT, J. R., and WILLIAMS, C. E. 1954. Trans. Inst. Radio Engrs. **AP-2**, 15.
DAY, J. W. B. and PSUTKA, M. E. 1956. Defence Research Telecommunications Establishment, Ottawa, Canada, Report No. 22-0-12.
FRIEND, A. W. 1945. Proc. Inst. Radio Engrs. **33**, 358.
GODSON, W. L. 1951. Quart. J. Roy. Meteorol. Soc. **77**, 633.
GORDON, W. E. 1955. Proc. Inst. Radio Engrs. **43**, 23.
GRANT, H. D. 1944. Cloud and weather atlas (Coward-McCann, Inc., New York).
HAY, D. R., DAY, J. W. B., and MAYNARD, L. A. 1958. Defence Research Telecommunications Establishment, Ottawa, Canada, Report No. 22-0-13.
HAY, D. R. and POAPS, G. E. 1959. Defence Research Telecommunications Establishment, Ottawa, Canada, Report No. 1013.
HIRAO, K., UYESUGI, Y., and TAO, K. 1952. J. Geomag. Geoelect. **4**, 131.
KLINKER, L. 1956. Hochfr. Tech. u. Electroak. **65**, 77.
LAUTER, E. A. and KLINKER, L. 1954. Nachrtech. Z. **4**, 242.
NORTHOVER, F. H. 1955. Can. J. Phys. **33**, 316.
PALMEN, E. and NEWTON, C. W. 1948. J. Meteorol. **5**, 220.
PENNER, C. M. 1955. Quart. J. Roy. Meteorol. Soc. **81**, 89.
PICKARD, G. W. and STETSON, H. T. 1947. Proc. Inst. Radio Engrs. **35**, 1445.
RANDALL, D. L. 1954. Bull. Am. Meteorol. Soc. **35**, 56.
REED, R. J. 1957. Bull. Am. Meteorol. Soc. **38**, 357.
SAWYER, J. S. 1955. London Meteorological Office, Geophys. Mem. **12** (96).

LIMITING PRECISION IN OPTICAL INTERFEROMETRY[1]

G. R. HANES

ABSTRACT

The fundamental limit to the precision of setting on a fringe peak due to photon noise is evaluated for a certain class of interferometric methods in terms of parameters characterizing the source, the interferometer, and the detector. It is shown that changes in path difference of 10^{-10} should be detectable with an observing time of 1 second, using only modest equipment. Some of the experimental conditions required to attain this precision are discussed.

I. INTRODUCTION

The corpuscular nature of radiant energy imposes a limit to the precision* of any optical measurement that occupies a finite observing time. Thus, even in an ideal experiment in which all of the available photons are effectively used or "counted", there will be a statistical spread in the results of successive measurements of the same quantity, that is, the precision is finite. It is possible to calculate the maximum theoretical precision expected of any particular method of measurement, and a comparison of this with the observed precision will indicate how successful the experimenter has been in eliminating at least the fluctuating part of possible perturbing effects that can influence the results. The steady part, or bias, of such effects reduces the accuracy* but not the precision of the measurement. In principle these perturbations can be eliminated, leaving only the fundamental quantum limitation to precision.

Some of the most careful interferometric measurements in recent years have been made in connection with an international program that has culminated in a resolution, to be presented to the International Conference on Weights and Measures, redefining the Metre in terms of the vacuum wavelength of the 6056 Å line of Kr^{86}. The wavelength of this radiation has been determined in terms of the cadmium red line at a number of national laboratories. The over-all accuracy of this comparison, as judged by the agreement of the results obtained at the various laboratories, is about 3×10^{7}. The precision reported for some of the individual results was somewhat higher.

The question arises as to how close we have come to the theoretical limit of precision in measurements of this type. As a partial answer to this question, a certain class of interferometric methods using modern techniques is described, and analyzed in detail in Part II. The maximum precision that can be expected is evaluated in terms of parameters characterizing the source, the interferometer, and the detector. A number of conclusions are drawn from the functional form of the expression, and numerical examples are given. Since the

[1]Manuscript received July 15, 1959.
Contribution from the Division of Applied Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5385.

*In this discussion, precision is defined as the reciprocal of the relative standard deviation of nominally identical measurements. Accuracy is taken to be a measure of the reciprocal of the relative spread of results obtained by a number of independent observers.

theoretical precision turns out to be much higher than the accuracy obtained at present, Part III is devoted to a consideration of some of the experimental conditions that must be established to attain this precision.

The general form of the apparatus to be considered is sketched in Fig. 1. A source S is imaged by lenses L_1 and L_2 in the plane of an aperture A, centered

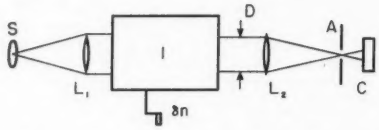


FIG. 1. Schematic diagram of experimental arrangement.

on the optic axis. Light from each point of S passes through the interferometer I as a parallel beam and emerges from I with a diameter D. The interferometer, which may be a Fabry–Perot etalon, or a two-beam-type like the Michelson, has its reflecting surfaces adjusted to be normal to the axis and therefore produces fringes of equal inclination centered on A. The interferometer also has an arrangement for making small adjustments δn in the order of interference, by making small changes in the optical path difference.

If the normal optical path difference in the interferometer is $2t$, the order of interference at the center of the ring system at A has the value

$$n = 2t\sigma \tag{1}$$

where σ is the wave number at the center of the spectral line. (It is assumed that the line is symmetric.) As t is changed, the intensity of the light passed by a very small aperture at A is shown in Fig. 2 as a function of n. To determine

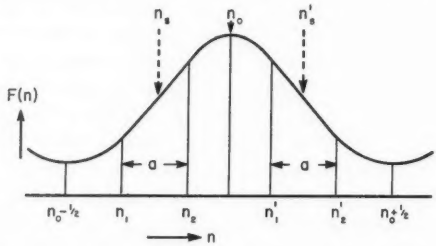


FIG. 2. Variation of flux with order of interference when a small aperture is used.

n_i corresponding to an initial value of $t = t_i$, t is reduced by a measured amount to t_0 such that n becomes an integer, as at n_0 in Fig. 2, where $F(n)$ has a maximum, that is, we set on the fringe peak. The fractional part of n_i follows immediately from the change in t necessary to make the setting; it is $2(t_i - t_0)\sigma$. The integral part of n_i can be obtained by several methods and will not be considered here.

The practical method of setting on the fringe peak is to compare the intensities $F(n_s)$, $F(n_s')$ at order numbers n_s, n_s' which are located symmetrically

about n and separated by a fixed interval. The criterion for making the setting is now that $F(n_s) = F(n_s')$, for it is evident that in this case $n = n_0$. It will be seen in the detailed analysis below that, in practice, the aperture at A should not be very small, but has an optimum size. According to our earlier discussion, the required intensities cannot be measured exactly, and thus there will be a finite minimum departure δn_m from the condition $n = n_0$ that can be detected above the unavoidable noise level at the detector. The maximum precision with which it is possible to determine n is then taken to be $n/\delta n_m$.

An instrument using essentially the above technique with a Fabry–Perot etalon was described by Baird (1954). The setting was made by visual detection of minimum flicker with a 10 c.p.s. scan between n_s and n_s' obtained by varying the air pressure with a motor-driven piston. A later version with photoelectric detection was used to measure the shift of some mercury lines with carrier gas pressure (Baird and Smith 1957). Terrien (1958*b*) has applied the method to a Michelson interferometer, making static readings of photomultiplier current first at n_s and then at n_s'. He has calculated the precision to be expected of his instrument when the limiting sources of noise were (*a*) shot effect in the photocathode current and (*b*) lamp fluctuations. In the analysis which follows it is considered that only the former is a fundamental limitation, the lamp fluctuations being in principle eliminable. On this basis, the limiting precision will be evaluated for both two-beam and multiple-beam interferometers with plane reflecting surfaces, and obtained in a form that displays its dependence on the parameters of the apparatus.

II. ANALYSIS

Each point in the plane of A (Fig. 1) is characterized by its angular displacement ϕ from the optic axis as seen from the center of L_2. At an angle ϕ_ϵ, the order of interference will be less than that at the center n_c, by an amount ϵ, where

$$n_c - \epsilon = 2t\sigma \cos \phi_\epsilon = 2t\sigma(1 - \phi_\epsilon^2/2). \tag{2}$$

From (1) and (2) we obtain immediately

$$\phi_\epsilon^2 = \epsilon/t\sigma.$$

If L_1 has a focal length f, the area of the lamp used to illuminate ϵ orders is

$$\pi f^2 \phi_\epsilon^2 = \pi f^2 \epsilon/t\sigma. \tag{3}$$

The solid angle over which light is collected from each point in the source is

$$\pi D^2/4f^2. \tag{4}$$

If the source has a radiance P (watts/metre^2 steradian) in the light of a given line, the power incident on the interferometer which is available to produce the interference pattern for ϵ orders is, from (3) and (4),

$$P \frac{\pi f^2 \epsilon}{t\sigma} \frac{\pi D^2}{4f^2} = \frac{\pi^2 D^2}{4t\sigma} P\epsilon .$$

Since this is linear in ϵ, the power available for each order is the same, and the power emerging from the interferometer in each order is

$$p = \frac{\pi^2 D^2}{4t\sigma} P\mathscr{T} \tag{5}$$

where $\mathscr{T}$ is the transmittance of the instrument averaged over one order. For the Michelson interferometer

$$\mathscr{T} = 2T_S R_S R_M$$

where T_S, R_S are the transmittance and reflectance of the beam splitter, and R_M is the reflectance of the mirrors.

In a practical instrument $\mathscr{T}$ will be somewhat less than its maximum theoretical value of 0.5, but the assumption of this value will not significantly overestimate the performance. For the Fabry–Perot etalon

$$\mathscr{T} = T^2/1-R^2$$

where T, R are the transmittance and reflectance of each etalon plate. Because of the multiple reflections, the effect of absorption in the reflecting films is more important than in the Michelson, and the actual values of T and R must be used in the above expression to obtain $\mathscr{T}$. The power p (equation 5) is distributed over the order as $pF(n)dn$ where $F(n)$ is normalized such that

$$\int_{n_0-\frac{1}{2}}^{n_0+\frac{1}{2}} F(n)dn = 1.$$

To make a setting on the fringe peak in an observing time τ the fluxes received from each side of the fringe during $\tau/2$ are compared. Since the fluctuations in these two signals are independent, the root-mean-square fluctuation in their difference will be the same as that in the sum, i.e. the noise will be determined by the total number of photons M_a received during time τ. M_a is given by

$$M_a = \frac{p\tau}{E}\int_{n_1}^{n_2} F(n)dn \tag{6}$$

where $E = hc\sigma$ is the energy of one photon and n_1, n_2 are the order numbers at the edges of the band passed by the aperture A (Fig. 1). (We need only consider the signal from one side of the fringe since the signals from the two sides are equal when the setting has been made.) With only a small error it may be assumed that if the source is steady, photons are emitted randomly, and thus the r.m.s. fluctuation in M_a will be $M_a^{\frac{1}{2}}$. In the presence of this noise we take the smallest detectable departure of the setting from the peak, δn_m, to be that value of δn which produces a difference signal M_s just equal to the r.m.s. fluctuation $M_a^{\frac{1}{2}}$, where

$$M_s = \frac{p\tau}{E}\left\{F(n_1) - F(n_2)\right\}\delta n. \tag{7}$$

Of the photons that fall on the detector, only a fraction θ produce a countable output. (θ has been called the quantum yield (Rose 1948) and the detective quantum efficiency (Jones 1958). It will be equal to the quantum efficiency of the primary photon detection mechanism only if there are no appreciable noise sources between the detector and the final indicator.) To calculate δn_m we must now equate θM_s to $(\theta M_a)^{\frac{1}{2}}$. Using equations (1), (5), (6), and (7), and assuming a spectral line width $\Delta\sigma$, the maximum precision with which n can be determined is

$$\frac{n}{\delta n_m} = \left[\frac{\pi^2}{2hc} D^2 \tau\theta \frac{P}{\Delta\sigma} Q\right]^{1/2} \tag{8}$$

where Q is given by

$$Q = \mu \mathscr{V} \frac{\{F(n_1)-F(n_2)\}^2}{\int_{n_1}^{n_2} F(n)dn}$$

and μ is the fraction of an order occupied by the line width $\Delta\sigma$, i.e. $\mu = 2t\Delta\sigma$.

The next problem is to find values of n_1 and n_2 that make Q a maximum. A Gaussian line shape will be assumed, but the general results will not be sensitive to line shape. For the Michelson interferometer

$$F(n) = 1+V\cos 2\pi n$$

where V is the visibility of the fringes, given by

$$V = \exp(-\pi^2\mu^2/4 \ln 2).$$

For the Fabry–Perot etalon

$$F(n) = 1+2\sum_{k=1}^{\infty} R^k \exp(-\pi^2\mu^2k^2/4 \ln 2)\cos 2\pi kn.$$

This expression was derived by Krebs and Sauer (1953).

Numerical investigation of several representative curves $F(n)$ revealed that Q was largest when the interval $a(=n_2-n_1)$ embraced most of the side of the fringe, and that a value of Q within 10 to 20% of the maximum could be obtained in all cases by choosing n_1 and n_2 such that

$$\begin{aligned} F(n_2)-F(n_0+1/2) &= 0.8\{F(n_0)-F(n_0+1/2)\}, \\ F(n_1)-F(n_0+1/2) &= 0.2\{F(n_0)-F(n_0+1/2)\}. \end{aligned} \tag{9}$$

These assumptions greatly simplify the calculation of Q, and retain enough accuracy for our purpose. It follows immediately from (9) that

$$F(n_2)-F(n_1) = 0.6\{F(n_0)-F(n_0+1/2)\}.$$

If $F(n)$ is approximated by a straight line between n_1 and n_2,

$$\int_{n_1}^{n_2} F(n)dn = \frac{1}{2}a\{F(n_0)+F(n_0+1/2)\}. \tag{10}$$

Then

$$Q = 0.72\mathscr{T}\frac{\mu}{a}\frac{\{F(n_0)-F(n_0+1/2)\}^2}{F(n_0)+F(n_0+1/2)}.$$

For the Michelson interferometer, a simple calculation shows that $a = 0.205$. Since $F(n_0) = 1+V$ and $F(n_0+1/2) = 1-V$,

$$\frac{\{F(n_0)-F(n_0+1/2)\}^2}{F(n_0)+F(n_0+1/2)} = 2V^2 = 2\exp(-\pi^2\mu^2/2\ln 2).$$

Taking $\mathscr{T} = 1/2$, Q becomes

$$Q = 3.51\,\mu\exp(-\pi^2\mu^2/2\ln 2).$$

The optimum $\mu = (\ln 2)^{1/2}/\pi = 0.265$ leads to a maximum value of $Q = 0.565$.

The calculations for the Fabry–Perot etalon could not be done as readily because of the complication of the function $F(n)$. Numerical values were determined from the tabulated curves given by Krebs and Sauer (1953), and are shown in Fig. 3 as a function of μ for several values of R, for dielectric and

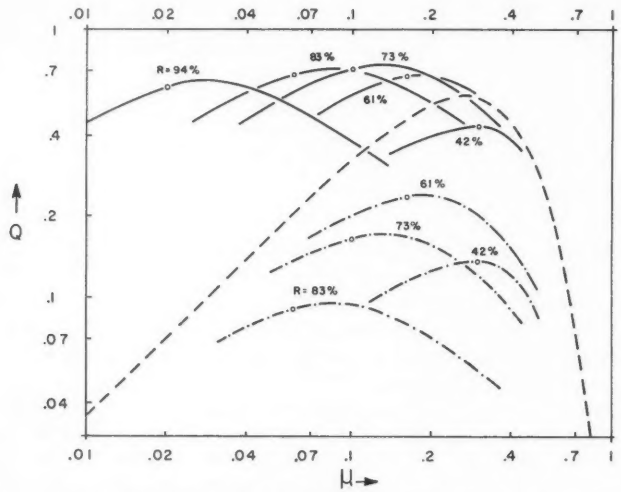


FIG. 3. Variation of Q with μ. Fabry–Perot: —— dielectric films, ·–·– aluminum films; Michelson: – – –.

aluminum reflecting films. The individual curves for aluminum are the same shape as those for dielectric films with the same value of R, but the heights are reduced because of the higher absorption. This effect is particularly severe for large R. The circles on these curves indicate the point at which the line width is equal to the instrumental width. The curve for the Michelson interferometer is also shown on the same figure.

It is seen from Fig. 3 that the maximum attainable precision is about the same for the Michelson interferometer and the Fabry–Perot, unless it is necessary to use aluminum films. (Aluminum has very little dispersion of phase change on reflection, a property that is useful when comparing wavelengths in different parts of the spectrum.) In this case, the Fabry–Perot suffers more because of the multiple reflections.

The path difference or etalon length at which the optimum precision is obtained is relatively short, t ranging from about 1/14 to 1/7 of $(\Delta\sigma)^{-1}$ depending on the instrument. The smallest $\Delta\sigma$ obtained from conventional sources is given by the Kr^{86} lamp operated at the triple point of nitrogen. For this source t ranges from about 5 to 10 cm. The optimum reflectance to be used in the Fabry–Perot is surprisingly low, about 70% for dielectric and 55% for aluminum films.

An important fact displayed in equation (8) is that the precision attainable depends on the characteristics of the spectral line only as $P/\Delta\sigma$. (There is an indirect effect arising from the wavelength dependence of θ for the detector used.) For a given line shape, $P/\Delta\sigma$ is directly related to the power emitted per unit spectral interval at the center of the line. Thus any arrangement for producing a narrower line, such as cooling of the source, or the use of an atomic beam, will lead to a loss of attainable precision unless the peak intensity is maintained. In practice, a smaller value of $\Delta\sigma/\sigma$ is highly desirable, since the interferometer is used at larger path difference where both n and δn_m are larger for the same precision. The larger δn_m allows a relaxation of the tolerances on steadiness of the source, uniformity of illumination, flatness, and parallelism of the interferometer mirrors and, of course, larger fringe fractions are easier to measure. These practical gains may lead to an acceptance of a source with a smaller limiting precision (i.e. $P/\Delta\sigma$) if its smaller half width allows a higher precision in actual experimental practice. It should be noted that accepting a loss of only a factor of 2 in the theoretical maximum precision allows an increase of μ to approximately 0.6. The same practical advantages accompany this change as those resulting from the use of a narrower line.

As an example, the value of equation (8) will be calculated for the 6056 Å line of a Kr^{86} discharge lamp operated at the triple point of nitrogen. MKS units are used throughout. For this line $P = 3.0\times10^{-1}$ (Engelhard 1958), $\Delta\sigma = 1.4$ (Terrien 1958*a*), with a current density in the capillary of 2×10^{-9} (2 ma/mm²). A photomultiplier with a trialkali cathode, such as the TA series of EMI, has $\theta = 7.0\times10^{-2}$ at this wavelength. Using an etalon with a 5-cm spacer and reflecting plates 4 cm in diameter coated with dielectric films to give $R = 73\%$ leads to the following: $D = 4.0\times10^{-2}$, $\mu = 1.4\times10^{-1}$, $Q = 7.4\times10^{-1}$. With an observing time $\tau = 1$ second, (8) becomes

$$n/\delta n_m = 2.1\times10^{10}.$$

For the green line of a Megger's type Hg^{198} lamp, the value of $\theta P/\Delta\sigma$ is about five times larger with the same photomultiplier. The maximum precision for this line is therefore about 4.7×10^{10}. The value of the precision obtainable with the krypton line confirms the estimate made by Terrien (1958*b*) for his

Michelson interferometer. His calculated precision was 10^9 for a photocathode current of about 10^5 electrons per second. The value used here (see below, Section III*c*) is 10^8 electrons per second, which gives a precision about 30 times greater, i.e. about 3×10^{10}.

An experimental confirmation of (8) is afforded by some figures kindly supplied by D. S. Smith. These results were obtained while measuring the shift of the krypton orange line (Baird and Smith 1959) using a later version of the apparatus described by Baird and Smith (1957). Putting in the constants relating to his apparatus, equation (8) gives $n/\delta n_m = 6\times10^8$. In terms of a detectable wavelength shift this is 10 μÅ. The estimated precision of a single setting was about 20 μÅ. Since a shift just equal to the r.m.s. fluctuations would not be reliably detected these values are considered to be in good agreement.

III. EXPERIMENTAL CONDITIONS

In this section some of the experimental conditions that must be established to allow an approach to the above theoretical precision are evaluated. Numbers used for illustration relate to the first example described above using the krypton orange line. The material in this part is intended to give some idea of the problems involved in attaining such high precision. It should be stressed that the attainment of an accuracy of measurement of the same order is a vastly more difficult task, since there are very many sources of systematic, but usually constant, errors that are not even mentioned here.

(*a*) *The Fringe Fraction*

The analysis in Part II shows that a change in fringe fraction of δn_m from the setting on the peak produces a noticeable output from the detector. (It is assumed, of course, that sufficient amplification is provided between the detector and the final indicator to show the detector noise.) In the example, $n = 1.6\times10^5$, and therefore $\delta n_m = 8\times10^{-6}$. To make use of this sensitivity, the mechanism for adjusting the fringe fraction must be fine enough that changes of this size can be made and read. This requirement is not prohibitive since many experiments require only a small range of adjustment of n.

(*b*) *Stability of the Interferometer*

To attain the precision in the example, changes in t other than those deliberately introduced to make the setting must be less than 5 parts in 10^{11} (i.e. 1/40 Å in 5 cm) over any period of time taken for measurements. This requires a high order of mechanical stability, and the temperature must be held constant to about 10^{-4} °C even with invar or fused quartz construction.

(*c*) *Lamp Stability*

A straightforward method of setting on the fringe peak is to measure the flux from one side of the fringe for a time $\tau/2$, shift to the other side of the fringe for $\tau/2$, and compare the results. It is evident that the difference in average flux leaving the source in two successive intervals of $\tau/2$ must be less than the variation in the number of counted events that are expected from

quantum fluctuations. The relative r.m.s. fluctuation in output due to quantum effects is $(\theta M_s)^{-1/2}$, where M_s is given by (6).

The assumptions (9) and (10) led to a value of the integral in (6) of 0.205 for the Michelson. The same value, within 5%, was found in all the cases considered for the Fabry–Perot as well. Thus M_s may be calculated from

$$M_s = 0.21 \frac{p\tau}{E}.$$

In the example above, $M_s = 1.4\times10^9$ and $(\theta M_s)^{-1/2} = 1.0\times10^{-4}$. Thus the stability of the light output from the source, as described above, must be better than 0.01%.

It might be mentioned at this point that the photocathode current in this example, $i = \theta M_s e = 1.6\times10^{-11}$ amperes, is much larger than the dark current ($<10^{-14}$ amperes). Thus the shot noise in the dark current is negligible and, ignoring the small amount of noise introduced by the electron multiplier, the quantum efficiency of the photocathode may be taken for the detective quantum efficiency of the tube.

Another method of setting on the fringe has been used at the National Research Council (Baird and Smith 1957). In this method the order is changed relatively rapidly (15 cycles per second) between n_s and n_s' in a sinusoidal scan, the mean value of n being adjusted to give a null in the output of a synchronous detector operated at the scanning frequency. The lamp fluctuations that affect the setting precision are in a narrow frequency band at the scanning frequency. This method will then be less sensitive to lamp fluctuations if the spectrum of these fluctuations has a decreasing quadratic content at higher frequencies.

A brief experiment to investigate this point was performed. Light from the source being studied fell on the cathode of a photomultiplier tube. The current from the anode passed through a load resistor and a microammeter. The alternating voltage across the load was amplified, passed through a narrow band filter tuned to 16 c.p.s., and read on an r.m.s. meter. From previous measurements on this phototube, the r.m.s. meter reading as a function of average anode current was known for a steady source. This of course is the amplified shot noise in the photocathode current. If the lamp fluctuations are large enough, a higher meter reading will be expected. Since the lamp fluctuations can be considered as a modulation of the incoming light, the relative r.m.s. value will be independent of the photomultiplier current, whereas the relative shot noise will vary inversely as the square root of the current. By using the whole spectrum from the source, much larger currents than are normally encountered could be obtained, and therefore lamp fluctuations that would normally be buried in the shot noise could be measured.

For the Kr lamp operated as in the example above from an unregulated power supply, the relative r.m.s. fluctuation in a 1 c.p.s. pass band at 16 c.p.s. was 8×10^{-5}. Thus the fluctuations in output of this lamp will not appreciably reduce the precision when the rapid scanning technique is used. With the static method of setting described first, the important fluctuations are those

occurring in successive 1/2-second intervals. These are very much larger, approaching 1% in some cases, although the exact r.m.s. value was not determined. Considerable stabilization of the source would therefore be necessary in this case. Similar measurement made on an electrodeless Hg lamp excited from an unregulated oscillator at about 300 Mc/s showed relative r.m.s. fluctuations of about 1.6×10^{-5} in a 1 c.p.s. band at 16 c.p.s.

(*d*) *Illumination*

Lack of perfect flatness and parallelism of the reflecting surfaces in the interferometer will cause variations of t over the field. The effective value of t will then depend on the illumination of the field.

The center of the aperture (A, Fig. 1) corresponds to n_2 and n_2' (Fig. 2), and since these are not symmetrically located relative to the fringe peak at n_0, changes in illumination over the aperture will cause an apparent shift of the peak. Approximate calculations of these two effects show that variation of the order of 10^{-3} in the illumination may lead to errors of a few parts in 10^{10} in certain types of measurements. However, in most experiments the relative illumination will be constant, and no loss of precision will occur.

IV. CONCLUSIONS

The principal results of this analysis are given in equation (8) and Fig. 3. They show that the precision obtainable with the Fabry–Perot and the Michelson interferometer are of the same order, and confirm a previous calculation for a Michelson interferometer with photoelectric detection (Terrien 1958*b*). It is shown that in contrast to the requirements for high resolution, optimum precision of setting on a fringe is obtained with a relatively short path difference, a relatively large aperture in the plane of the Haidinger rings, and, for the Fabry–Perot, mirrors of quite low reflectance. According to the theory, the only important property of the source is $P/\Delta\sigma$, the radiance in the light of a given line divided by the line width. However, in practice a small line width gives important experimental advantages. An additional consideration in practice is the wavelength dependence of θ, the detective quantum efficiency of available detectors.

ACKNOWLEDGMENTS

The author is indebted to Dr. K. M. Baird and Mr. D. S. Smith for many stimulating discussions. He is also grateful to Dr. K. H. Hart for useful comments on the manuscript, and to Mr. W. E. E. Berger for assistance in the measurement of lamp fluctuations.

REFERENCES

Baird, K. M. 1954. J. Opt. Soc. Am. **44**, 11.
Baird, K. M. and Smith, D. S. 1957. Can. J. Phys. **35**, 455.
——— 1959. Can. J. Phys. (in press).
Engelhard, E. 1958. Procès-Verbaux des Séances du CIPM 2me série, B, **26**, M66.
Jones, R. C. 1958. Phot. Sci. Eng. **2**, 57.
Krebs, K. and Sauer, A. 1953. Ann. Physik, **13**, 359.
Rose, A. 1948. Advances in Electronics, **1**, 131.
Terrien, J. 1958*a*. Procès-Verbaux des Séances du CIPM 2me série, B, **26**, M135.
——— 1958*b*. J. phys. radium, **19**, 390.

THE SPECTRUM OF Li^+ [1]

G. HERZBERG AND H. R. MOORE[2]

ABSTRACT

The spectrum of the Li^+ ion has been reinvestigated under high resolution in the region 10,000 to 1000 Å with a view to determining the Lamb shifts in the ground state and first-excited state. The hyperfine structure of the line 5485 Å has been studied under improved resolution both with the natural isotopic mixture and with a 96% Li^6 sample. For many of the higher levels the hyperfine structure is larger than the fine structure and for large n and l it is even larger than the singlet–triplet separation. The coupling conditions in these cases are discussed in some detail and theoretical intensity distributions in the combined fine structure and hyperfine structure are presented for comparison with the partially resolved observed structures. The selection rule $\Delta G = 0$ for the quantum number of the total spin $\boldsymbol{G} = \boldsymbol{I}+\boldsymbol{S}$ is found to be obeyed.

On the basis of this detailed analysis a table of energy levels of Li^+ is presented including, where possible, the hyperfine splittings. The absolute accuracy of the energy values is only $\pm 3\,cm^{-1}$ because of the error introduced by the resonance lines at $1\,^1S$–$n\,^1P$ below 200 Å not remeasured in the present work. However, the relative accuracy of the levels with reference to the $2\,^3S$ level is considered to be within $\pm 0.05\ cm^{-1}$ in most cases.

The limits of the observed series have been determined with an accuracy of $\pm 0.1_0\ cm^{-1}$ and lead to an ionization potential of

$$I.P._{exp}(Li^+) = 610079._4\ cm^{-1}$$

with an error ($\pm 3\ cm^{-1}$) caused entirely by the error in the resonance lines. The term values (ionization potentials) of the $2\,^3S$ and $2\,^1S$ states

$$T(2\,^3S) = 134044.19 \pm 0.10\ cm^{-1}$$
$$T(2\,^1S) = 120008.30 \pm 0.10\ cm^{-1}$$

are much more accurate. The difference of the observed ionization potential of Li^+ and that calculated from relativistic quantum mechanics is $-8.0 \pm 3\ cm^{-1}$, which agrees within the rather large limit of error with the predicted Lamb shift ($-7.8\ cm^{-1}$).

A. INTRODUCTION

Most of the experimental and theoretical work on the Lamb shift refers to the one-electron systems H, D, He^+. However, the problem of Lamb shifts in two-electron systems has been attacked recently. Kabir and Salpeter (1957) and Sucher (1958) have carried out extensive calculations leading to a predicted value for the shift of the ground state of He and to estimates for Li^+ and higher ions of this series. Experimentally, the Lamb shift can be determined as the difference of the observed ionization potential and the ionization potential calculated on the basis of the Dirac theory (neglecting quantum-electrodynamic effects). Both quantities have now been obtained with considerable precision in the case of He (Herzberg 1958; Pekeris 1958) and their difference agrees in a very satisfactory way with the predicted Lamb shift.

[1]Manuscript received August 13, 1959.
Contribution from the Division of Pure Physics, National Research Council, Ottawa, Canada.
Issued as N.R.C. No. 5391.
[2]National Research Council Postdoctorate Fellow, 1955–57. Present address: Electro-Optical Systems, Inc., 170 N. Daisy Ave., Pasadena, California.

Since the accuracy of neither the predicted nor the observed values for the Lamb shift of He is very high, it appears desirable to obtain experimental Lamb shifts in a few other cases of energy levels of two-electron systems. For this reason, work on the spectrum of Li^+ was initiated in this laboratory.

An experimental determination of the Lamb shift is possible not only for the ground state but also for the metastable $2\,^3S$ state of two-electron systems. A theoretical (Dirac) value for the energy of the $2\,^3S$ level of He has been calculated by Pekeris (1959) and that for Li^+ is being calculated by Hart and Herzberg (1959). For He a precise experimental value of the energy of the $2\,^3S$ level was implied in the earlier work on the ground state of He, namely

$$T(2\,^3S) = T(2\,^3P_1) + \nu(2\,^3S\text{–}2\,^3P_1) = 29223.86 + 9230.87 = 38454.73 \pm 0.05\ \text{cm}^{-1}.$$

Previous work on the spectrum of Li^+ has been summarized by Moore (1949). Term values were obtainable at that time only within 5 cm^{-1} and the ionization potential was estimated to have an error of ± 25 cm^{-1}. Such a low accuracy is clearly insufficient to determine Lamb shifts and for this reason the present measurements were undertaken. It has been possible to obtain the term values of a large number of excited states (including $2\,^3S$) with an accuracy of ± 0.10 cm^{-1} but a definitive term value for the ground state (i.e. for the ionization potential) has not yet been obtained, although a considerable improvement over the previous accuracy has been accomplished. For a precise determination of the ionization potential the resonance lines at 199.3, 178.0, . . . Å will have to be remeasured with a much increased accuracy. Unfortunately these lines are in a spectral region that we have not yet reached with our present equipment.

B. LIGHT SOURCE

Because of the high ionization and excitation potential of Li^+ (the highest of any singly ionized system) its spectrum is not easy to obtain, particularly since the low vapor pressure and high chemical activity of Li add further difficulties. This explains the scarcity of previous work on the spectrum of Li^+ and the low accuracy obtained.

Three different kinds of sources were tried in the present experiments, (*a*) a high-frequency, high-voltage discharge with external electrodes, (*b*) a sliding spark (see Bockasten 1955), and (*c*) a hollow cathode tube. The high-frequency discharge did not yield any Li^+ lines. The sliding spark gave a good intensity of Li^+, but apart from other difficulties the lines were quite broad, presumably on account of Stark effects due to interionic fields. The hollow cathode was by far the best source both for its intensity and for the narrowness of the lines produced and it was therefore chosen for the present investigation. The original design of Schüler (1926) was slightly modified. The cylindrical cathode of 6.5-cm length and 1.4-cm diameter was made of molybdenum, which was found to withstand the chemical activity of lithium at high temperatures fairly well. The internal diameter of the cathode was 1.1 cm. One end was closed while the other had an opening of 0.6-cm diameter. These dimensions approximate those recommended by Risberg (1956) in his work on Na I and

PLATE I

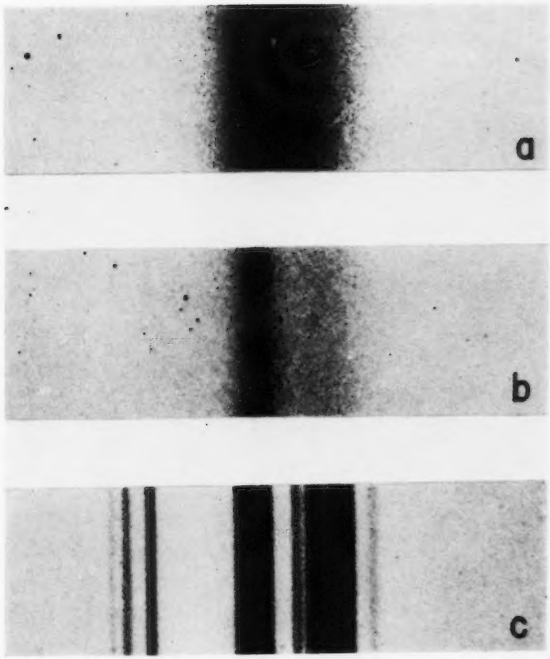


FIG. 1. Effect of excitation conditions on the fine structure of the 5485 Å line of Li^{++}. (*a*) Sliding spark at 14 kv, with a series inductance of 1 turn. (*b*) Sliding spark at 8 kv, with a series inductance of 30 turns. (*c*) Hollow cathode, at 250 v and 300 ma.

PLATE II

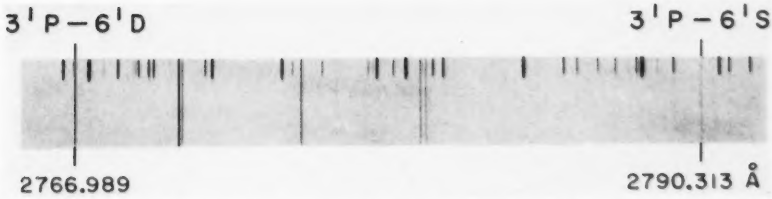


FIG. 2. Li^{+} spectrum in the region 2760–2800 Å.

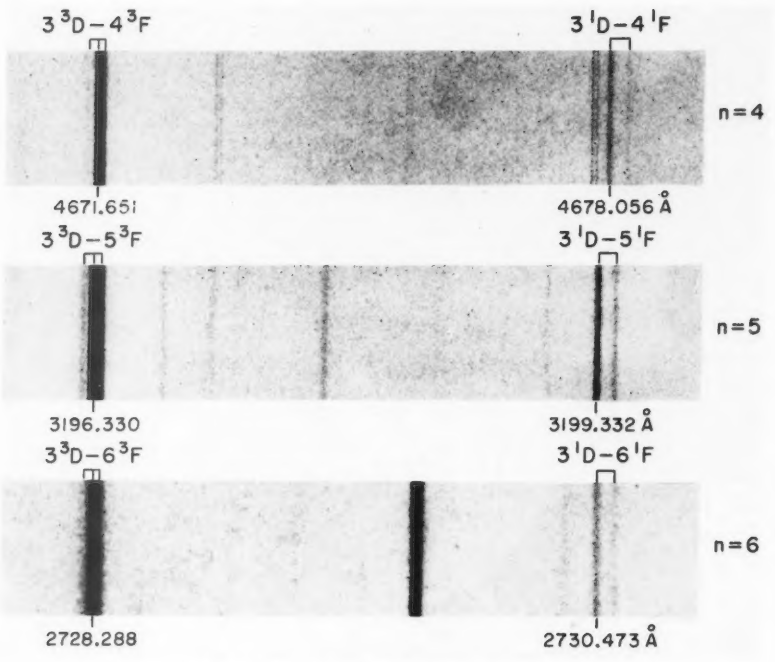


FIG. 3. Lines of the Bergman series of Li^+.

K I. The molybdenum cylinder was surrounded by a close-fitting copper jacket of external diameter 2.5 cm. This jacket ensured a more uniform distribution of temperature along the cathode than one would otherwise obtain. The cathode with its axis horizontal was supported by a vertical thin-walled steel tube which insulated the cathode thermally and at the same time, when necessary, could be used to cool the cathode by a stream of cold air. In front of the opening of the cathode was a small iron ring which served as the anode. Although at the same (ground) potential as the whole metal vacuum chamber, this anode appeared to stabilize the discharge and to lead to an increased intensity of the Li^+ lines.

For the operation of the discharge, helium at a pressure of 1 mm was used as a carrier gas. A d-c. potential of 350 volts was applied and currents up to 300 ma were used. A charge of about 0.3 g of lithium was introduced into the cathode and was sufficient for exposures up to 8 hours. The cathode could be heated up to 1000° C directly by the discharge. A separate furnace, as used by Schüler, was not found necessary to evaporate the lithium.

The superiority of the hollow cathode source over the sliding spark source is shown by Fig. 1 which reproduces spectra obtained with the two sources of the fine structure of the Li^+ line 5485 Å.

C. OBSERVED SPECTRA AND STANDARDS

The spectra in the photographic infrared, visible, and near ultraviolet regions (8600–2000 Å) were taken in the first or second orders of the N.R.C. 21-ft and 35-ft grating spectrographs. The spectra in the vacuum region (2000–1180 Å) were taken in the fourth and fifth orders of the 3-meter vacuum spectrograph.

Above 2000 Å, Fe and Ne lines from a separate hollow cathode tube, photographed in the same order as the Li^+ lines, were used as standards. In order to eliminate any possibility of shifts between the Li^+ lines and the comparison spectrum (on account of changes of temperature or barometric pressure), the latter was exposed partly before and partly after the main exposure. Figures 2 and 3 show sections of these spectra.

In the vacuum ultraviolet, in order to obtain comparable wave number accuracy, it is imperative to obtain higher wavelength accuracy than in the visible and near ultraviolet regions. Therefore, internal standards occurring in the same source as the spectrum to be measured and in the same wavelength region are desirable. Such internal standards near the five Li^+ lines in the vacuum region are lines of Ca II and Mg II derived from the combination principle by Edlén and Risberg (1956) and Risberg (1955) and the resonance lines of N I near 1200 Å measured by Herzberg (1958). The former were produced in the Li^+ source by adding small amounts of calcium or magnesium to the hollow cathode, while the latter were obtained by mixing a trace of nitrogen with the carrier gas. As an example, Fig. 4 shows the Li^+ lines 1493.0 and 1653.1 Å together with two Ca^+ lines. Other Ca^+ lines in the second spectrogram, as well as Mg^+ lines in the first, are beyond the region reproduced. The spectra shown were taken in the fourth order of the

vacuum spectrograph and were separated from the overlapping second and third orders by the use of a foregrating as described by Douglas and Herzberg (1957). A spectrum of Fe from the hollow cathode tube mentioned is superimposed in second order (by suitable rotation of the foregrating) but this serves only for the determination of the shape of the correction curve whose absolute position is determined by the Ca^+, Mg^+, and N lines.

The exposure times varied considerably depending on the lines studied and the experimental conditions. To obtain the strongest components of the

TABLE I

Wavelengths and wave numbers of Li^+ lines[c] (except λ 5485 Å, see Table II)

Transition	Present work λ, Å	I_{rel}[a]	ν_{vac}, cm^{-1}	Previous work λ, Å Werner	Schüler
$2\,^1S$–$2\,^1P$	8517.369		11737.49	(7.1)[b]	
$3\,^1P$–$4\,^1S$	5037.915		19843.95	7.8	8.2
$3\,^3P$–$4\,^3S$	4881.490	1	20479.83		
	4881.386	4	20480.27	1.3	
	4881.320	4	20480.54		
$3\,^1P$–$4\,^1D$	4788.363		20878.13	8.8	8.5
$3\,^1D$–$4F$	4678.290	1	21369.35		
	4678.056	3	21370.42	8.48	8.2
$3\,^3D$–$4F$	4671.705	2	21399.47		
	4671.651	6	21399.72	1.66	
	4671.509	0	21400.37		
?	4607.338		21698.43	7.4	
$3\,^3P$–$4\,^3D$	4325.514	1	23111.98		
	4325.471	5	23112.37	5.42	
	4325.419	5	23112.65		
$3\,^1S$–$4\,^1P$	4516.447		24052.23	6.3	
?	3714.580	0	26913.29		
	3714.509	1	26913.80		
	3714.414	10	26914.49	4.1	4.41
	3714.400	7sh.	26914.59		
	3714.292	8	26915.37		4.27
	3714.273	6sh.	26915.51		
	3714.162	5	26916.31		4.18
	3714.004	1	26917.46		
$3\,^3S$–$4\,^3P$	3684.320		27134.32	4.1	
$3\,^1P$–$5\,^1S$	3306.284		30236.73	5.2	6.3
$3\,^1P$–$5\,^1D$	3249.868		30761.61	9.8	9.9
$3\,^1D$–$5F$	3199.434	2	31246.49		
	3199.332	5	31247.49	9.4	9.43
$3\,^3D$–$5F$	3196.356	4	31276.59		
	3196.330	9	31276.84	6.3	
	3196.264	1	31277.49		
$3\,^3P$–$5\,^3S$	3155.330	4	31683.23	5.4	
	3155.308	3	31683.45		
$3\,^3P$–$5\,^3D$	3029.136	3	33003.10	9.1	
	3029.121	3	33003.27		
?	2934.251	1	34070.28		4.26
	2934.115	5	34071.86	4.2	4.13
	2934.066	2	34072.43		4.07
	2934.022	2	34072.94		
$3\,^1P$–$6\,^1S$	2790.313		35827.71		0.39
$3\,^1P$–$6\,^1D$	2766.989		36129.70	7.0	7.0
$3\,^1D$–$6F$	2730.551	1	36611.81		
	2730.473	3	36612.86	0.7	0.5
$3\,^3D$–$6F$	2728.315	2	36641.81		
	2728.288	5	36642.17	8.4	
	2728.239	0	36642.83		
$3\,^3S$–$5\,^3P$	2674.460		37379.61	4.4	
$3\,^3P$–$6\,^3S$	2657.303	3	37620.94	7.3	
	2657.293	2	37621.08		
$3\,^3P$–$6\,^3D$	2605.081		38375.06	5.1	
$3\,^1P$–$7\,^1D$	2539.487		39366.20		9.4
$3\,^1D$–$7F$	2508.785		39847.92	8.9	8.83
$3\,^3D$–$7F$	2506.940		39877.26	7.0	
$3\,^3P$–$7\,^3S$	2429.814		41142.93	30.0	
$3\,^1P$–$8\,^1D$	2410.842		41466.68		0.85
$3\,^3P$–$7\,^3D$	2402.331		41613.57	2.3	
$3\,^1D$–$8F$	2383.199		41947.61		3.26
$3\,^1D$–$8F$	2381.538		41976.88	1.6	
?	2336.997	2	42776.85		
	2336.910	5	42778.44	7.0	
	2336.875	3	42779.08		
$3\,^3S$–$6\,^3P$	2329.836		42908.31	30.0	
$3\,^3P$–$8\,^3S$	2303.330		43402.04		
$3\,^3D$–$9F$	2302.568		43416.40		
$3\,^3P$–$8\,^3D$	2286.822		43715.33		
$3\,^3D$–$10F$	2249.211	†	44446.26		
$2\,^1P$–$3\,^1S$	1755.331$_7$		56969.29	5.4	
$2\,^1P$–$3\,^1D$	1681.660$_7$		59465.03	1.8	
$2\,^3P_0$–$3\,^3S$	1653.212$_9$	1	60488.28		
$2\,^3P_2$–$3\,^3S$	1653.132$_0$	5	60491.24	3.3	
$2\,^3P_1$–$3\,^3S$	1653.076$_7$	3	60493.26		
$2\,^3P_0$–$3\,^3D$	1493.036$_8$	1	66977.60		
$2\,^3P_2$–$3\,^3D$	1492.973$_0$	5	66980.45	3.1	
$2\,^3P_1$–$3\,^3D$	1492.931$_1$	3	66982.33		
$2\,^3S$–$3\,^3P$	1198.092$_4$		83466.01	8.0	

NOTE: †Diffuse line; sh. "shoulder", incompletely resolved weaker companion of a stronger line.

[a]The estimated intensities in this column are relative intensities within each multiplet. The numbers given for different multiplets are not comparable.

[b]Series and Willis (1958).

[c]Note added in proof: In his thesis, which has only now become available to us, Werner (1927) gives the following additional lines: $2\,^1S$–$3\,^1P$ 1420.7 Å, $2\,^1P$–$4\,^1S$ 1253.8 Å, $2\,^3P$–$4\,^3S$ 1166.4 Å, and $2\,^3P$–$4\,^3D$ 1132.1 Å.

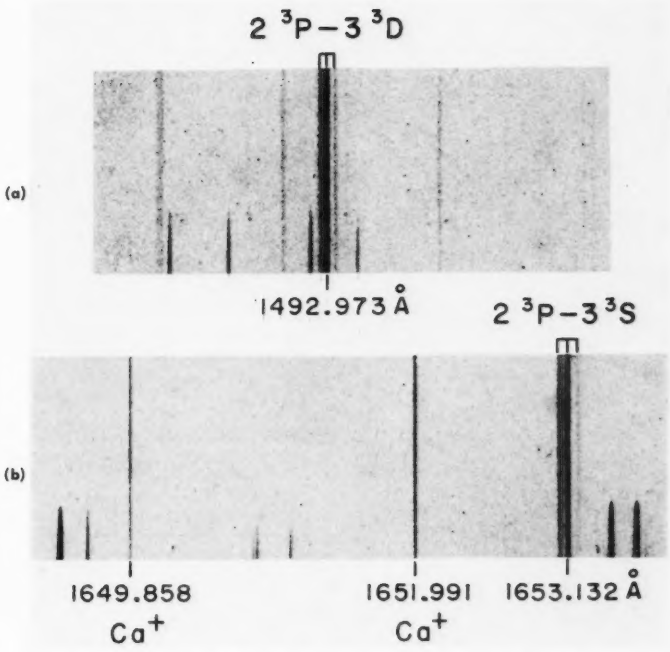


FIG. 4. Sections of the Li^+ spectrum in the vacuum ultraviolet.

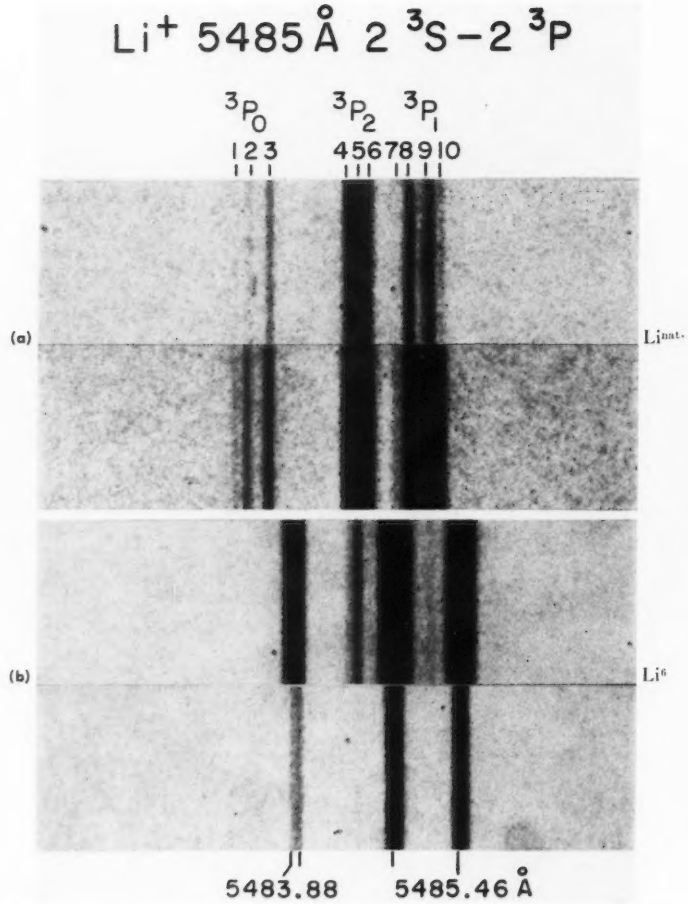


FIG. 5. Structure of the Li^+ line 5485 Å; (*a*) for natural lithium, and (*b*) for lithium in which Li^6 was enriched to 96%. In each case two exposures of different density are given. The components are numbered as in Table II.

line 5485 Å in the second order of the 21-ft grating only 30 seconds were required, while some of the higher members of the series required up to 10 hours.

The plates were measured by means of a photoelectric comparator similar to that described by Tomkins and Fred (1951) and built by B. P. Stoicheff in this laboratory. Most lines were measured on several plates. The average wavelengths and wave numbers of all observed Li^+ lines except the line 5485 Å are given in Table I and compared with the previous measurements of Werner (1925, 1926) and Schüler (1924, 1925, 1926). The infrared line was only recently observed under low dispersion by Series and Willis (1958). By comparison of measurements on different plates it is estimated that the measurements above 2000 Å are accurate to ±0.003 Å, while those below 2000 Å are accurate to ±0.001 Å. This corresponds throughout to a wave number accuracy of ±0.05 cm^{-1} or better. Estimated relative intensities are given only for triplets and apply only within each triplet.

The line 5485 Å is well known to have a somewhat complex structure (Schüler 1925, 1927, 1930; Granath 1932). Figure 5*a* shows this line as obtained in the second order of our 21-ft grating. This spectrum was obtained with natural lithium (92.5% Li^7, 7.5% Li^6) as were most of the other spectra. Figure 5*b* shows a spectrum obtained with a sample of Li in which the isotope Li^6 was enriched to 96%. Comparison of these two spectra makes it possible to sort out the Li^6 and Li^7 lines. These are listed separately in Table II and compared with the data of Schüler (1927). It may be noted that our wavelengths for the components of this line are all about 0.21 Å smaller than

TABLE II

Fine structure and hyperfine structure of the Li^+ line 5485 Å of Li^7 and Li^6

Component	number	λ air, obs. Present work	λ air, obs. Schüler	I_{est}	ν_{vac}, obs. (present work)	Triplet comp.	Hyperfine comp.		ν_{vac}, calc.	I_{calc}
Li^7	(1)	5483.345$_8$	3.55	1	18231.97$_5$		1/2–3/2	*a*	18231.97$_5$	0.520
	(2)	5483.450$_0$	3.67	2	18231.59$_8$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_0$	3/2–3/2	*b*	18231.59$_2$	1.179
	(3)	5483.655$_9$	3.86$_4$	4	18230.94$_4$		5/2–3/2	*c*	18230.94$_9$	2.337
							1/2–3/2	*d*	18228.51$_0$	1.191
	(4)	5484.424$_4$	4.63	6	18228.38$_2$		3/2–5/2	*e*	18228.45$_7$	3.287
							1/2–1/2	*f*	18228.29$_4$	1.541
	(5)	5484.485$_2$	4.69	10	18228.18$_7$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_2$	5/2–7/2	*g*	18228.21$_5$	8.000
							3/2–3/2	*h*	18228.12$_4$	2.448
	(6)	5484.595$_9$	4.78$_5$	5	18227.81$_9$		3/2–1/2	*i*	18227.90$_8$	0.459
							5/2–5/2	*j*	18227.81$_4$	2.713
		5484.697$_9$		1	18227.480		5/2–3/2	*k*	18227.48$_1$	0.361
							1/2–3/2	*l*	18226.57$_4$	2.289
	(8)	5484.973$_6$	5.17$_5$	8	18226.56$_4$		3/2–5/2	*m*	18226.52$_9$	2.713
							1/2–1/2	*n*	18226.43$_6$	0.459
	(13)	5485.134$_0$	5.31	5	18226.03$_1$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_1$	3/2–3/2	*o*	18226.18$_6$	0.374
							3/2–1/2	*p*	18226.05$_0$	1.541
	(9)	5485.176$_7$	5.37$_0$	7	18225.88$_9$		5/2–5/2	*q*	18225.88$_6$	3.287
	(10)	5485.280$_5$	5.47	3	18225.54$_4$		5/2–3/2	*r*	18225.54$_2$	1.337
Li^6	(14*a*)	5483.879$_8$	4.10	3	18230.19$_3$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_0$	1–1	*s*	18230.19$_7$	
	(14*b*)	5483.920$_0$		4	18230.06$_6$		2–1	*t*	18230.06$_7$	
	(7)	5484.823$_8$	5.02	6	18227.06$_3$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_2$		*u*	18227.04$_6$	
	(11)	5485.463$_9$	5.65	5	18224.93$_5$	$2\,^2S_{\frac{1}{2}}$–$2\,^2P_1$		*v*	18224.95$_3$	

Schüler's. This shift is due to the fact that Schüler measured in second order against the Li line at 2741.2 Å appearing in fourth order assuming an old value 2741.29 for this line which more recently (France 1930; Datta and Bose 1935; Johansson 1959) has been revised to 2741.184 Å. This change accounts for the difference. Our wavelengths, based on measurements of five different plates, are believed to supersede those of Schüler.

D. TERM SYSTEM, FINE STRUCTURE, AND HYPERFINE STRUCTURE

The assignment of most of the observed lines has already been made by Werner (1925, 1926) and Schüler (1924, 1926). In Fig. 6 an energy level diagram is plotted showing the observed lines. The only unassigned lines or multiplets are λ 3714.2 and 2934.2 already observed by Schüler (1925) and Werner (1926) and 4607.4 and 2337.0 Å previously observed by Werner (1926). These lines do not result from transitions between levels shown in the diagram Fig. 6. It is difficult to be certain that they are due to Li^+ although they behave in the same way as the Li^+ lines. If they are due to Li^+ they must be ascribed to transitions between doubly excited states.

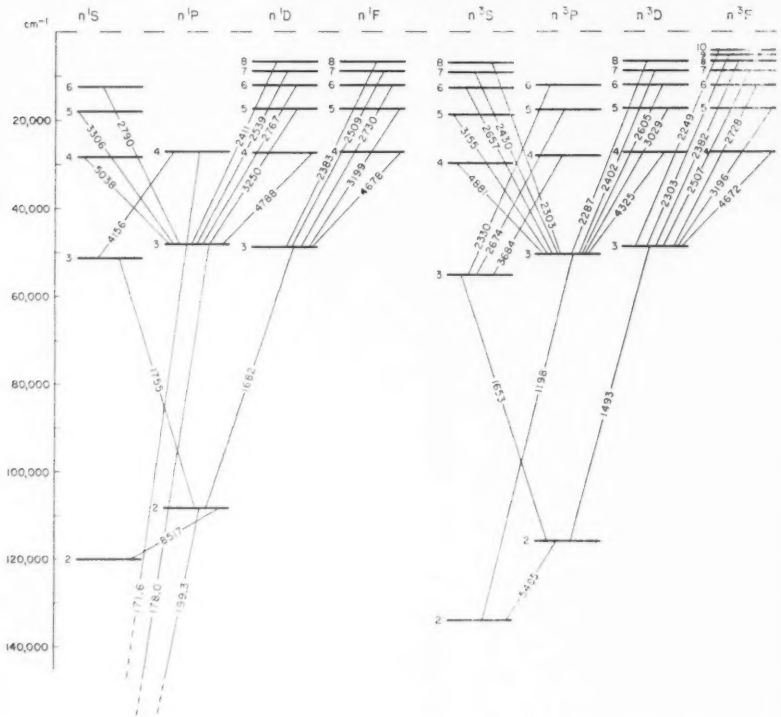


FIG. 6. Energy level diagram of Li^+. All observed levels except the $1\,^1S$ ground state are shown and the observed transitions are indicated. (See, however, footnote *c* of Table I.)

With the exception of the lines $2\,^3S$–$2\,^3P$ and $2\,^1S$–$2\,^1P$ all the classified lines observed in the visible and near ultraviolet regions have as their lower states three-quantum states. All transitions between the three-quantum and two-quantum states lie in the vacuum region.

The detailed analysis of the Li^+ spectrum and particularly the determination of the series limits is made complicated by the fact that the hyperfine structure splittings are of the same order as or even larger than the ordinary fine structure splittings. This complication has been well known for many years for the line 5485 Å corresponding to the transition $2\,^3S$–$2\,^3P$. But this is the only line whose fine structure had previously been resolved (Schüler 1925, 1927, 1930). As is well known the nuclear spins of Li^7 and Li^6 are $\frac{3}{2}$ and 1 respectively. The present analysis is based on these values and in a way may be said to confirm them if that were necessary.

As shown by Table II, a slightly better resolution has been obtained for the 5485 Å line than by Schüler. His assignment of the Li^6 lines has been confirmed by the observation of the spectrum of nearly pure Li^6 (Fig. 5*b*). In this spectrum, component (14) has been partially resolved into two components which must correspond to the two stronger of the three expected components (0–1, 1–1, 2–1). Indeed the separation of these two components agrees well with that predicted from the known ratio of the magnetic moments of Li^7 and Li^6 [$\mu(Li^7)/\mu(Li^6) = 3.9614$, see Kopfermann (1958)] and the observed splitting factor ($A = 0.257$) of the $2\,^3S_1$–$2\,^3P_0$ component of Li^7. This predicted value is 0.130 while 0.13_3 is observed. The lines (7) and (11) of Li^6 are not further resolved since the hyperfine splitting of $2\,^3P_1$ and $2\,^3P_2$ is of the same order as that of $2\,^3S_1$ and there is a greater number of line components within the same wave number interval.

Taking account of the statistical weights, the center of the $2\,^3S_1$ state of Li^6 is the ordinary mean of the levels $F = 2$ and $F = 1$, and therefore the mean of the components (14*a*) and (14*b*) when subtracted from components (7) and (11) gives the positions of $2\,^3P_2$ and $2\,^3P_1$ below $2\,^3P_0$. One finds 3.07_1 and 5.19_8 cm^{-1}, i.e., as is well known, the $2\,^3P$ state is partially inverted (see Bethe and Salpeter 1957). With the aid of these triplet intervals we can now obtain a calculated set of hyperfine levels both for Li^7 and Li^6. For Li^6 the standard hyperfine structure formulae (see Kopfermann 1958) can be used since the hyperfine splitting is small compared to the triplet splitting. However for Li^7, since the hyperfine splitting is much larger, the more elaborate method of Güttinger and Pauli (1931) has to be used which takes account of the perturbations between levels of the same F in $2\,^3P$.

Güttinger and Pauli have used the same splitting factor a for both $2\,^3S$ and $2\,^3P$. While indeed the main contribution to the splitting is made by the 1*s* electron, that is, a is the same in $2\,^3S$ and $2\,^3P$, the 2*s* electron in $2\,^3S$ would be expected to make a small contribution (calculated to be 0.019 from formula (26.9) of Kopfermann) while the contribution of the 2*p* electron should be entirely negligible. Therefore the a factor for $2\,^3P$ is expected to be somewhat smaller than for $2\,^3S$. Accepting the above predicted contribution 0.019 of the 2*s* electron, one would expect $a = 0.495$ for $2\,^3P$ if $a = 0.514$ for $2\,^3S$ (as

obtained from the three components 1, 2, 3). However, we find that $a = 0.495$ for $2\,^3P$ gives a slightly poorer representation of the observed components than $a = 0.514$. Thus we adopt (in agreement with Pauli and Güttinger) the same splitting factor for $2\,^3P$ and $2\,^3S$: $a = 0.514\ \text{cm}^{-1}$. The energy levels derived from this splitting factor and the observed line components are listed in Table III for both Li^7 and Li^6. The wave numbers of the lines derived

TABLE III

Energy levels $2\,^3S$ and $2\,^3P$ of Li^7 and Li^6 giving the best fit with the observed fine structure components in the line 5485 Å

Multiplet component	Li^7 F	Li^7 energy above center of $2\,^3S_1$, cm^{-1}	Li^6 F	Li^6 energy above center of $2\,^3S_1$, cm^{-1}
$2\,^3S_1$	1/2	−0.643	0	−0.130
	3/2	−0.257	1	−0.065
	5/2	+0.386	2	+0.065
$2\,^3P_1$	1/2	18225.793	0	18224.888
	3/2	18225.931	1	18224.920
	5/2	18226.272	2	18224.986
$2\,^3P_2$	1/2	18227.651	1	18226.948
	3/2	18227.867	2	18227.014
	5/2	18228.200	3	18227.111
	7/2	18228.604		
$2\,^3P_0$	3/2	18231.335	1	18230.132

from the energy levels of Table III are included in the second last column of Table II. The agreement with the observed wave numbers, particularly of unblended components, is satisfactory. The revised triplet splittings are:

$$2\,^3P_0\text{–}2\,^3P_2 = 3.086, \qquad 2\,^3P_0\text{–}2\,^3P_1 = 5.179\ \text{cm}^{-1}.$$

The calculated intensities are given in the last column of Table II. These were obtained by the method of Güttinger and Pauli but without some of the approximations made by them. The agreement with experiment is slightly better than found by Güttinger and Pauli; for example, the intensity ratio of components (3) and (2) was observed by Schüler (1930) to be 1.95 ± 0.04. According to Güttinger and Pauli this ratio was to be 1.51 and this discrepancy was considered to be serious by Schüler. However, the revised calculations give a ratio of 1.98, in excellent agreement with experiment.

None of the other lines of Li^+ has been resolved as completely but some resolution has been obtained for all triplet transitions and in a few cases even for singlets. If there were no hyperfine structure one would expect only the transitions $2\,^3P$–$3\,^3S$, $2\,^3P$–$3\,^3D$ and $3\,^3P$–$n\,^3S$, $3\,^3P$–$n\,^3D$ to be split while $3\,^3S$–$n\,^3P$ ($n > 4$) and $3\,^3D$–$n\,^3F$ should be single because the triplet splitting for $n\,^3P$ with $n > 4$ and for all ^3D and ^3F states would be expected to be below $0.2\ \text{cm}^{-1}$, i.e. unobservable with the resolving power of our gratings. Actually, as is shown by Fig. 3, all the lines of $3\,^3D$–$n\,^3F$ do show a splitting which could *formally* be explained as due to a splitting of $3\,^3D$ but it is five times as large as the expected triplet splitting. In addition, each of the singlet

lines $3\,^1D$–$n\,^1F$ has a satellite of appreciable intensity even though on the basis of ordinary hyperfine structure theory no splitting of singlet levels is expected.

In order to understand these apparent anomalies it is necessary to remember that the hyperfine splitting in Li^+ and similar two-electron systems is determined mainly by the $1s$ electron, present in all levels considered, and that therefore the splitting in all triplet levels has about the same magnitude as in the $2\,^3S$ state (see Table III). Even the singlet levels will show such a splitting if their separation from the corresponding triplet levels (of equal n) is smaller than the hyperfine splitting of the latter. Unfortunately it is not possible to use standard hyperfine structure formulae since the coupling conditions are quite different from the usual ones. While for small hyperfine splitting the coupling may be written symbolically

$$[(\boldsymbol{l}_1\boldsymbol{l}_2)(\boldsymbol{s}_1\boldsymbol{s}_2)]\boldsymbol{I} = (\boldsymbol{LS})\boldsymbol{I} = \boldsymbol{JI} = \boldsymbol{F}$$

for large hyperfine splitting and small triplet splitting we have (with $\boldsymbol{l}_1 = 0$)

$$\text{(I)} \qquad [(\boldsymbol{s}_1\boldsymbol{s}_2)\boldsymbol{I}]\boldsymbol{l}_2 = (\boldsymbol{SI})\boldsymbol{l}_2 = \boldsymbol{Gl}_2 = \boldsymbol{F}.$$

It may even happen that the hyperfine splitting is large compared with the singlet–triplet energy difference. In this case the coupling may be described by

$$\text{(II)} \qquad (\boldsymbol{s}_1\boldsymbol{I})\boldsymbol{s}_2\boldsymbol{l}_2 = (\boldsymbol{F}_i\boldsymbol{s}_2)\boldsymbol{l}_2 = \boldsymbol{Gl}_2 = \boldsymbol{F}.$$

In case (I) $\boldsymbol{J}$ loses its meaning. Instead there is a quantum number G corresponding to the resultant of $\boldsymbol{S}$ and $\boldsymbol{I}$ to which $\boldsymbol{l}_2$ is very loosely coupled. In case (II) the nuclear spin $\boldsymbol{I}$ couples with $\boldsymbol{s}_1$ (or $\boldsymbol{j}_1$, if $l_1 \neq 0$) to form a resultant $\boldsymbol{F}_i$, the $\boldsymbol{F}$ value of the ion; $\boldsymbol{F}_i$ is very slightly coupled with $\boldsymbol{s}_2$ and $\boldsymbol{l}_2$. As before $\boldsymbol{J}$ is not defined.

Case (I) can be treated by the Güttinger–Pauli method which takes into account the interaction of levels of the same F. By applying this method to a state with small triplet splitting, one finds indeed that the resulting levels are arranged in three groups corresponding to $G = 1/2, 3/2, 5/2$ and for each G value there are $2l_2+1$ sublevels if $l_2 < G$ and $2G+1$ if $l_2 > G$. This is shown for $3\,^3D$ and $4\,^3F$ in Fig. 7(a). (Such an arrangement can also be seen in the $3\,^3D$ levels of He^3 discussed by Fred, Tomkins, Brody, and Hamermesh (1951)). The splitting between levels of the same G decreases rapidly with increasing n and l.

Since G is a pure spin quantum number (electron spin + nuclear spin) one would expect the selection rule $\Delta G = 0$ to hold. As a result the transition between two states belonging to case (I) should consist of one strong line ($\Delta G = 0$), possibly with weak satellites (with $\Delta G \neq 0$). The observed structure of the $3\,^3D$–$n\,^3F$ lines (Table I) might have been understood on this basis: there is a strong central line ($\Delta G = 0$), a very weak short-wave satellite at 0.64 cm^{-1} which would fit the transition from $G = 5/2$ to $G = 3/2$ (i.e. $\Delta G = +1$), and a much stronger and closer long-wave satellite at 0.25 cm^{-1} which would correspond to the splitting of the sublevels with $G = 3/2$ in the lower state, e.g. the transition from $F = 7/2$ to $F = 5/2$ (see Fig. 7(a)) with $\Delta G = 0$.

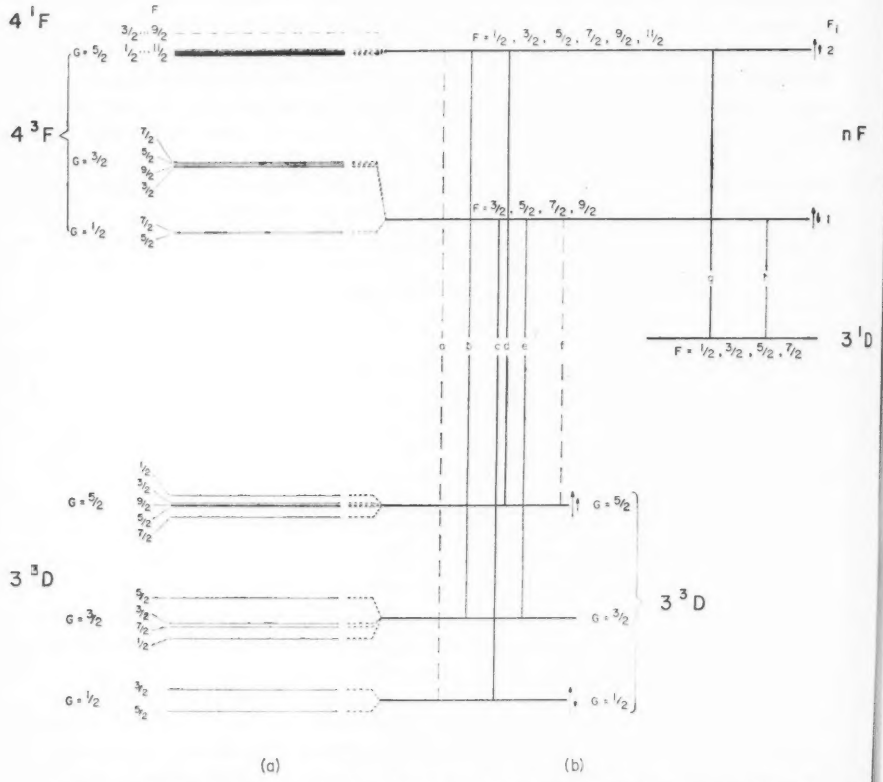


FIG. 7. Energy level diagram for $3\,^3D$–$4\,^3F$, $3\,^1D$–$4\,^1F$, theoretical: (*a*) in coupling case (I), (*b*) using case (I) for ^3D and case (II) for $4\,^3F$ and $4\,^1F$, disregarding all smaller interactions.

However, this interpretation cannot be maintained because an additional complication arises from the fact that the $n\,^3F$ levels have very nearly the same energy as $n\,^1F$, that is, that the $n\,^3F$–$n\,^1F$ energy difference is much smaller than the hyperfine structure splitting. Therefore coupling scheme (II) must be used for the F states. In a first approximation the splitting is determined by that in the ground state of the Li^{++} ion which is 0.995 cm^{-1} (calculated from the known magnetic moments of Li^7 and H and the observed splitting in hydrogen by means of standard formulae, see Kopfermann (1958)). Thus for large n we expect two component levels with $F_i = 1$ and 2 separated by approximately 1 cm^{-1}. With decreasing n each of these levels would be expected to split into several component levels (6 for $F_i = 1$, and 10 for $F_i = 2$, since $j_2 = 2\frac{1}{2}$ and $3\frac{1}{2}$) whose spacing will be small compared to 1 cm^{-1} even for the lowest F levels $4\,^3F$, $4\,^1F$. Since S is not defined the two main levels ($F_i = 1, 2$) are, so to speak, mixtures of ^3F and ^1F and therefore one

would expect both of them to combine with the $3\,^1D$ state. It is now very significant that the stronger members of the $3\,^1D$–$n\,^1F$ series are indeed observed to be double consisting of one strong and one weak component which are separated by 1.07, 1.00, and 1.05 cm^{-1} for $n = 4, 5$, and 6 respectively (see Fig. 3). Within the accuracy of the measurements this splitting approaches the predicted value 0.995 cm^{-1} of the Li^{++} ground state. It must be noted that the stronger component is at the shortward side, that is, corresponds to the upper component level with $F_i = 2$.

A very similar anomaly has been observed previously for the $4\,^1F$–$n\,^1G$ series of Al^+ by Paschen (1923) and Sawyer and Paschen (1927), and has been interpreted by Paschen (1932), Goudsmit and Bacher (1933), Paschen and Ritschl (1933), and more recently and in greater detail by Suwa (1950, 1951). The latter authors have also discussed the effect of this peculiar coupling on the corresponding triplet transition $4\,^3F$–$n\,^3G$. Similar transitions in In^+ have been observed and discussed by Paschen (1935) and Paschen and Campbell (1938).

While in Al^+ and In^+ the lower state $4\,^3F$ shows a fairly normal triplet structure, for Li^+ the lower state of the triplet transitions has a very small triplet splitting which is completely overshadowed by the hyperfine splitting leading to coupling scheme (I). Thus we have a transition between a state ($n\,^1F$, ^3F) belonging to coupling scheme (II) and a state ($3\,^3D$) belonging to coupling scheme (I). Rigorous intensity rules for such a case have not been worked out but we may attempt to apply the same correspondence considerations that were used by Goudsmit and Bacher, namely that those transitions in which the relative orientation of the spin of the $1s$ electron and of the nuclear spin remains unchanged are the strongest. In Fig. 7(*b*) an energy level diagram for the $3\,^3D$–nF transition is shown neglecting all the weaker interactions, i.e. there are only two levels in the upper state ($F_i = 1, 2$) and three in the lower state ($G = 1/2, 3/2, 5/2$). In addition the $3\,^1D$ level is also shown. The range of F values is indicated for each level. The selection rule for F allows six transitions, a, b, c, d, e, f, for $3\,^3D$–nF and two, g and h, for $3\,^1D$–nF. In the levels $F_i = 2$ of the upper state and $G = 5/2$ of the lower state the spins $\boldsymbol{I}$ and $\boldsymbol{s}_1$ are parallel, in $F_i = 1$ and $G = 1/2$ they are antiparallel (see the arrows to the right of the levels). Therefore according to Goudsmit–Bacher's rule, c and d would be expected to be strongest, b and e intermediate, and a and f forbidden. As can be seen by comparing Fig. 7(*b*) with the spectrum, Fig. 3 and Table I, the strongest one of the three observed components can be interpreted as the superposition of the two transitions c and d while the other components can be interpreted as e and b.

Corresponding $n\,^3D$ and $n\,^1D$ states (and a fortiori $n\,^3P$ and $n\,^1P$) do not lie close enough to each other for coupling (II) to be considered except at very large n. Therefore all transitions other than $3\,^3D$–$n\,^3F$, $3\,^1D$–$n\,^1F$ can be dealt with on the basis of coupling scheme (I), that is, by an application of the Güttinger–Pauli method.

The fine structure and hyperfine structure of the transition $2\,^3S$–$3\,^3P$ can be predicted completely from a knowledge of the structure of $2\,^3S$–$2\,^3P$ by

means of elementary fine structure theory (see Bethe and Salpeter (1957)) and Güttinger and Pauli's method. The triplet splitting for $3\,^3P$ is predicted to be $^3P_0$–$^3P_2 = 0.90_4$, $^3P_2$–$^3P_1 = 0.62_2\,\mathrm{cm}^{-1}$ (for $4\,^3P$ the corresponding figures are 0.38_1 and $0.26_2\,\mathrm{cm}^{-1}$ respectively). On this basis, the relative positions of the component levels of $3\,^3P$ can be derived. In Fig. 8 the calculated structures of $2\,^3S$–$2\,^3P$, $2\,^3S$–$3\,^3P$, and $3\,^3S$–$4\,^3P$ are compared. Corresponding transitions are designated by the same letters (see Table II). It is seen that in accordance with the selection rule $\Delta G = 0$ which holds in coupling

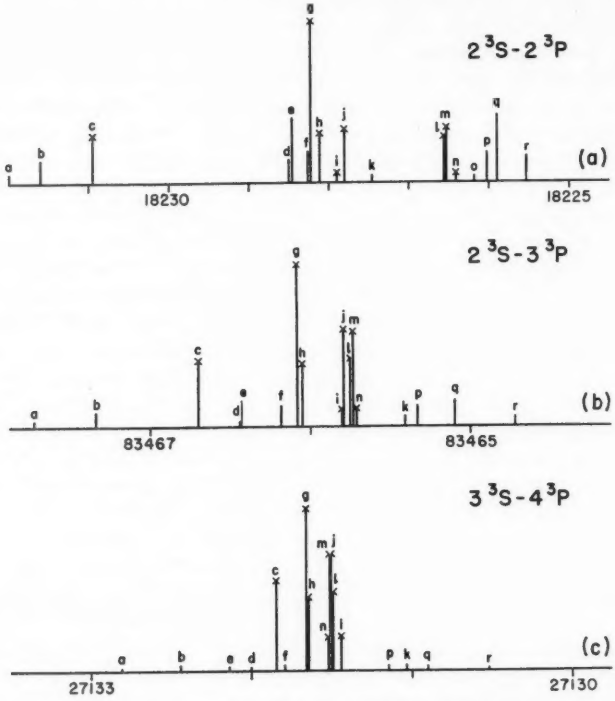


FIG. 8. Theoretical splitting patterns of (*a*) $2\,^3S$–$2\,^3P$, (*b*) $2\,^3S$–$3\,^3P$, (*c*) $3\,^3S$–$4\,^3P$. Note that the wave number scale in (*a*) is half that in (*b*) and (*c*). The lines marked × correspond to $\Delta G = 0$.

case (I), with increasing n the intensity is more and more concentrated in the transitions with $\Delta G = 0$ (marked by ×) while the transitions with $\Delta G \neq 0$ rapidly lose their intensity. At the same time the transitions with $\Delta G = 0$ move closer and closer together. With sufficient resolution the line $2\,^3S$–$3\,^3P$ would consist of three components with spacings 0.61 and $0.32\,\mathrm{cm}^{-1}$. However, in the vacuum ultraviolet such a resolution is difficult to obtain and in the present work only one line has been observed. From the position of this line the energies of the $3\,^3P$ levels can be derived as given in Table IV.

The near ultraviolet series $3\,^3S$–$n\,^3P$ is expected to be similar to $2\,^3S$–$n\,^3P$. For $3\,^3S$–$4\,^3P$ the structure represented in Fig. 8(*c*) is expected. Because of the smaller triplet splitting, the lines with $\Delta G \neq 0$ have still lower intensity than for $2\,^3S$–$3\,^3P$ and the strong lines with $\Delta G = 0$ are more concentrated near the center. In agreement with this prediction only a single line has been found. Even with the higher resolution in the near ultraviolet the close components near the center cannot be resolved. Similar considerations apply to $3\,^3S$–$5\,^3P$, and $3\,^3S$–$6\,^3P$.

The transition $2\,^3P$–$3\,^3S$ is expected to have a structure very similar to that of $2\,^3S$–$2\,^3P$ except for the slight change of hyperfine splitting of $3\,^3S$ compared to $2\,^3S$. Since this line is in the vacuum region complete resolution cannot be expected, but three components are indeed found as shown by Fig. 4(*b*). Their spacings are observed to be 2.02 and 2.96 cm^{-1}. These spacings agree well with the spacings of the centers of gravity of the corresponding predicted groups of lines (see Fig. 8(*a*)) which are 2.03 and 3.08 cm^{-1}.

Again the near ultraviolet series $3\,^3P$–$n\,^3S$ is similar to those just discussed. In fact the structure of the line $3\,^3P$–$4\,^3S$, except for the very slight change in the hyperfine splitting of $4\,^3S$ (and the reversal of the scale), should be identical with that of $2\,^3S$–$3\,^3P$ shown in Fig. 8(*b*). Three components have been observed whose spacings (0.44 and 0.27 cm^{-1}) agree reasonably well with the spacings of the centers of the three expected central line groups (0.63 and 0.34 cm^{-1}). The spacing is appreciably different from the predicted triplet spacings (0.904 and 0.622 cm^{-1}) used in the computation of Fig. 8(*b*). The strong effect of the perturbations in the hyperfine structure (affecting both position and intensity of the lines) is clearly shown in this way.

For the transition $2\,^3P$–$3\,^3D$ a rather complex structure is expected since the $2\,^3P$ state is still far removed from coupling case (I) while $3\,^3D$ is approaching it. The selection rule $\Delta J = 0, \pm 1$ is no longer valid and instead of the 45 transitions possible when the hyperfine structure is small compared to the fine structure, here all 56 transitions compatible with the rule $\Delta F = 0, \pm 1$ can occur. Figure 9(*a*) shows the intensity distribution derived from a detailed calculation according to the Güttinger–Pauli method. Unfortunately the available resolution in the vacuum region allows the observation of only three components (see Fig. 4 (*a*)) whose spacings (2.85 and 1.88 cm^{-1}) agree satisfactorily with those predicted from Fig. 9(*a*) by averaging the three line groups (2.94 and 1.86 cm^{-1}).

The transitions $3\,^3P$–$n\,^3D$ in the visible region are similar to $2\,^3P$–$3\,^3D$, but the lower state is closer to coupling case (I) and the upper state is an almost perfect case (I). Therefore the structure is somewhat simpler in that the transitions with $\Delta G = 0$ are more strongly emphasized compared to the other transitions. Figure 9(*b*) shows the calculated structure for $3\,^3P$–$4\,^3D$. In Fig. 10 the main transitions are shown in an energy level diagram. If the lower state were as good an approximation to case (I) as is the upper state, only one line would be observed which would be the superposition of the three transitions $1/2 \leftrightarrow 1/2$, $3/2 \leftrightarrow 3/2$, $5/2 \leftrightarrow 5/2$ in G corresponding to the group (*a*) of three lines at the left in Fig. 10. Because of the deviation from

FIG. 9. Theoretical splitting patterns of (*a*) $2\,^3P$–$3\,^3D$, (*b*) $3\,^3P$–$4\,^3D$. The wave number scale in (*a*) is half that in (*b*). Lines marked by × are in agreement with the selection rule $\Delta G = 0$. The broken lines represent the averages of the main line groups.

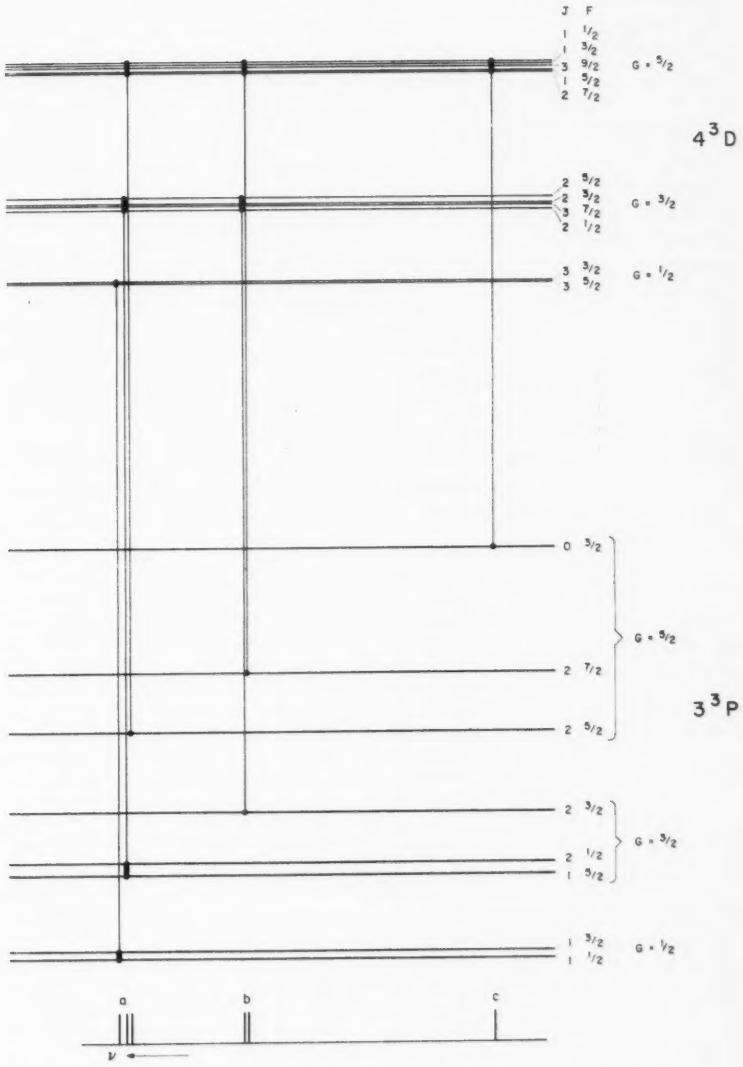


FIG. 10. Energy level diagram for $3\,^3P$–$4\,^3D$ showing principal transitions.

TABLE IV

Energy levels of $^7Li^+$

Designation	F	Energy
$1s^2\ ^1S_0$	$\frac{3}{2}$	$0(\pm 3)$
$2s\ ^3S_1$	$\frac{1}{2}$	476034.59_7
	$\frac{3}{2}$	476034.98_3
	$\frac{5}{2}$	476035.62_6
$2s\ ^1S_0$	$\frac{3}{2}$	490071.10
$2p\ ^3P_1^o$	$\frac{1}{2}$	494261.03_3
	$\frac{3}{2}$	494261.17_1
	$\frac{5}{2}$	494261.51_2
$2p\ ^3P_2^o$	$\frac{1}{2}$	494262.89_1
	$\frac{3}{2}$	494263.10_7
	$\frac{5}{2}$	494263.44_0
	$\frac{7}{2}$	494263.84_4
$2p\ ^3P_0^o$	$\frac{3}{2}$	494266.57_5
$2p\ ^1P_1^o$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	501808.59
$3s\ ^3S_1$	$\frac{1}{2}$	554754.08
	$\frac{3}{2}$	554754.45
	$\frac{5}{2}$	554755.06
$3s\ ^1S_0$	$\frac{3}{2}$	558777.88
$3p\ ^3P_1^o$	$\frac{1}{2}$	559500.31
	$\frac{3}{2}$	559500.35
	$\frac{5}{2}$	559500.72
$3p\ ^3P_2^o$	$\frac{1}{2}$	559500.78
	$\frac{3}{2}$	559501.03
	$\frac{5}{2}$	559501.42
	$\frac{7}{2}$	559501.71
$3p\ ^3P_0^o$	$\frac{3}{2}$	559502.32
$3d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	561243.15
$3d\ ^3D_{G=3/2}$	$\frac{1}{2}, \frac{3}{2}, \frac{7}{2}$	561243.60
	$\frac{5}{2}$	561243.77
$3d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	561244.30
$3d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	561273.62
$3p\ ^1P_1$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	561752.82
$4s\ ^3S_1$	$\frac{1}{2}$	579980.97
	$\frac{3}{2}$	579981.33
	$\frac{5}{2}$	579981.93
$4s\ ^1S_0$	$\frac{3}{2}$	581596.77
$4p\ ^3P^o_{G=1/2}$	$\frac{1}{2}, \frac{3}{2}$	581885.58
$4p\ ^3P^o_{G=3/2}$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	581885.98
$4p\ ^3P^o_{G=5/2}$	$\frac{3}{2}, \frac{5}{2}, \frac{7}{2}$	581886.70
$4d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	582613.02
$4d\ ^3D_{G=3/2}$	$\frac{1}{2}, \ldots, \frac{7}{2}$	582613.41
$4d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	582614.07
$4d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	582630.95
$4f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	582642.97
$4f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	582644.04
$4p\ ^1P_1^o$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	582830.11
$5s\ ^3S_1$	$\frac{1}{2}$	591183.90
	$\frac{3}{2}$	591184.26
	$\frac{5}{2}$	591184.86
$5s\ ^1S_0$	$\frac{3}{2}$	591989.55
$5p\ ^3P^o_{G=1/2}$	$\frac{1}{2}, \frac{3}{2}$	592133.65
$5p\ ^3P^o_{G=3/2}$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	592134.03
$5p\ ^3P^o_{G=5/2}$	$\frac{3}{2}, \frac{5}{2}, \frac{7}{2}$	592134.70
$5d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	592503.70
$5d\ ^3D_{G=3/2}$	$\frac{1}{2}, \ldots, \frac{7}{2}$	592504.09
$5d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	592504.75
$5d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	592514.43
$5f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	592520.11
$5f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	592521.11
$5p\ ^1P_1^o$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	592634.91
$6s\ ^3S_1$	$\frac{1}{2}$	597121.59
	$\frac{3}{2}$	597121.95
	$\frac{5}{2}$	597122.55
$6s\ ^1S_0$	$\frac{3}{2}$	597580.53
$6p\ ^3P^o_{G=1/2}$	$\frac{1}{2}, \frac{3}{2}$	597662.35
$6p\ ^3P^o_{G=3/2}$	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$	597662.73
$6p\ ^3P^o_{G=5/2}$	$\frac{3}{2}, \frac{5}{2}, \frac{7}{2}$	597663.40
$6d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	597875.55
$6d\ ^3D_{G=3/2}$	$\frac{1}{2}, \ldots, \frac{7}{2}$	597875.94
$6d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	597876.60
$6d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	597882.52
$6f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	597885.43
$6f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	597886.48
$7s\ ^3S_1$	$\frac{1}{2}$	600643.54
	$\frac{3}{2}$	600643.90
	$\frac{5}{2}$	600644.50
$7s\ ^1S_0$	$\frac{3}{2}$	$600930._6$
$7d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	601114.06
$7d\ ^3D_{G=3/2}$	$\frac{1}{2}, \ldots, \frac{7}{2}$	601114.45
$7d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	601115.11
$7d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	601119.02
$7f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	$601120._4$
$7f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	601121.55
$8s\ ^3S_1$	$\frac{1}{2}$	602902.62
	$\frac{3}{2}$	602902.98
	$\frac{5}{2}$	602903.58
$8d\ ^3D_{G=1/2}$	$\frac{3}{2}, \frac{5}{2}$	603215.82
$8d\ ^3D_{G=3/2}$	$\frac{1}{2}, \ldots, \frac{7}{2}$	603216.21
$8d\ ^3D_{G=5/2}$	$\frac{1}{2}, \ldots, \frac{9}{2}$	603216.87
$8d\ ^1D_2$	$\frac{1}{2}, \ldots, \frac{7}{2}$	603219.50
$8f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	$603220._3$
$8f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	603221.21
$9f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	604659.55
$9f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	604660.70
$10f\ ^3F^o\ F_i = 1$	$\frac{3}{2}, \ldots, \frac{9}{2}$	$605689._4$
$10f\ ^1F^o\ F_i = 2$	$\frac{1}{2}, \ldots, \frac{11}{2}$	$605690._5$

case (I) in the lower state a splitting arises even for the transitions with $\Delta G = 0$. The two "lines" (*b*) in Fig. 10 arise from $3\,^3P_{2,3/2}$ and $3\,^3P_{2,7/2}$ while the line (*c*) arises from $3\,^3P_{0,3/2}$. The three observed lines correspond apparently to the three groups (*a*), (*b*), and (*c*) with a few $\Delta G \neq 0$ transitions mixed in (see Fig. 9(*b*)). The observed spacings 0.28 and 0.39 agree satisfactorily with those predicted from Fig. 9(*b*), i.e. 0.30_8 and 0.51_1 cm^{-1} and the estimated intensities 5:5:1 are compatible with the predicted 15:11:5.

On the basis of the preceding discussion it is possible to establish the relative positions of the various energy levels of the triplet system. For the singlet system the situation is very much simpler and need not be discussed in detail. The relative position of the singlet and triplet system is determined by the singlet and triplet Bergman series since for both the upper states are the same, i.e. $F_i = 1$ and 2 (see Fig. 7). In Tabel IV the energy values of all observed levels above the $1\,^1S$ ground state are listed including, where possible, the individual hyperfine levels. In the case of the $n\,^3S$ and $2\,^3P$ and $3\,^3P$ levels, the calculated energy values are listed that give the best representation of the observed incompletely resolved hyperfine structures. The absolute accuracy of the energy values is only ± 3 cm^{-1} since they are based on approximate values of the far ultraviolet lines (see below) but the relative accuracy of the energy values for most excited states should be ± 0.05 cm^{-1} except that the $3\,^1P$, $4\,^1D$, $5\,^1D \ldots 4\,^1S$, $5\,^1S \ldots$ levels have a greater possible error relative to the other levels (perhaps ± 0.15 cm^{-1}) since they are related to them only by series limits.

E. SERIES LIMITS, TERM VALUES, AND IONIZATION POTENTIALS

If the usual methods of determining series limits (plotting n^*–n against T) are applied to the observed three singlet series and four triplet series (using the strongest lines in the latter) the limits listed in Table V are obtained. The

TABLE V

Series limits in Li^+

Series designation	Limit, cm^{-1}	n values used
$3\,^1D$–$n\,^1F$	48806.15±0.10	4, 5, 6, 7, 8
	48805.16±0.10[a]	4, 5, 6
$3\,^1P$–$n\,^1D$	48326.95±0.10	4, 5, 6, 7, 8
$3\,^1P$–$n\,^1S$	48327.20	4, 5, 6
$3\,^3D$–$n\,^3F$	48835.45±0.10[b]	4, 5, 6, 7, 8, 9, 10
$3\,^3P$–$n\,^3D$	50578.26±0.10[c]	4, 5, 6, 7, 8
$3\,^3P$–$n\,^3S$	50578.40±?	4, 5, 6, 7, 8
$3\,^3S$–$n\,^3P$	55326.50	4, 5, 6

[a] Weak satellite.
[b] Strongest component.
[c] Average of two strongest components.

two series $3\,^1P$–$n\,^1S$ and $3\,^1P$–$n\,^1D$, as expected, give very nearly the same limit, 48327.20 and 48326.95 cm^{-1} respectively. The latter value appears somewhat more reliable since the magnitude of the Rydberg correction (and therefore

also of the higher terms in the Ritz formula) is much smaller for $n\,^1D$ than for $n\,^1S$. We must, however, add a correction on account of the splitting of the ground state of Li^{++}. Since for all observed n values the ^1D and ^1S states are above the corresponding triplet states they must correlate with the upper ($F_i = 2$) component of the Li^{++} ground state. Referring to the center of gravity of the Li^{++} ground state, i.e. the energy that Li^{++} would have without the effect of nuclear spin, we have to subtract 3/8 of the splitting, that is 0.37 cm^{-1}, and therefore obtain for the term value of $3\,^1P$

$$T(3\,^1P) = 48326.58 \pm 0.10\ cm^{-1}.$$

Similarly the limit of the Bergman series (strong component) corresponds to the upper component of the Li^{++} ground state as discussed in detail in the preceding section. Correcting to the unsplit ground state of Li^{++} we obtain

$$T(3\,^1D) = 48805.78 \pm 0.10\ cm^{-1}$$

and combining with the wave numbers of the lines $2\,^1P$–$3\,^1D$ and $2\,^1S$–$2\,^1P$

$$T(2\,^1P) = 108270.81\ cm^{-1},$$

$$T(2\,^1S) = 120008.30\ cm^{-1}.^3$$

In deriving term values from the limits of the triplet series, the fine structure and hyperfine structure must be taken into account. According to the interpretation in Section D the observed limit of the strong component of the Bergman series ($3\,^3D$–nF) gives the term value of the $G = 5/2$ component of $3\,^3D$ below $F_i = 2$ or of the $G = 1/2$ component below $F_i = 1$ of the ion (see Fig. 7(*b*)). Referring again to the unsplit ground state of the Li^{++} ion we find

$$T(3\,^3D_{G=5/2}) = 48835.07\ cm^{-1},$$

$$T(3\,^3D_{G=1/2}) = 48836.08\ cm^{-1}.$$

The strongest observed component of the transition $2\,^3P$–$3\,^3D$ corresponds to $2\,^3P_{2,7/2}$–$3\,^3D_{3,9/2}$ where $3\,^3D_{3,9/2} \approx 3\,^3D_{G=5/2}$. Thus we obtain (see Table I)

$$T(2\,^3P_{2,7/2}) = 115815.5_2\ cm^{-1}.$$

The term values of the other components of $2\,^3P$ are readily obtained from this number and the energy values of Table III.

From the same table we obtain for the lowest triplet state (weighted average of hyperfine components) $T(2\,^3S_1) = 134044.1_2\ cm^{-1}$. A check on this term value can be obtained from the $3\,^3P$–$n\,^3D$ series. Only in the first two members of this series are the two components *a* and *b* of Fig. 10 resolved. The series limit corresponds to the average of these. The $G = 5/2$ group of levels of $n\,^3D$ must go to the $F_i = 2$ level of the ion and therefore according to Fig. 10 the series limit gives the average of the levels $3\,^3P_{2,7/2}$ and $3\,^3P_{2,5/2}$ below $F_i = 2$ of the ion. It also gives the average of $3\,^3P_{2,3/2}$, $3\,^3P_{2,1/2}$, $3\,^3P_{1,5/2}$ below $F_i = 1$. Adding the average energy of $3\,^3P_{2,7/2}$ and $3\,^3P_{2,5/2}$ above $2\,^3S$ on the one hand and that of $3\,^3P_{2,3/2}$, $3\,^3P_{2,1/2}$, $3\,^3P_{1,5/2}$ on the other

[3]See, however, "Note added in proof" at the end of this paper.

(Table IV) and correcting to the unsplit ground state of the ion yields the values $T(2\,^3S_1) = 134044.24$ and $134044.48\,cm^{-1}$. The mean of these two values, $134044.3_6\,cm^{-1}$, agrees satisfactorily with the value from the Bergman series but is perhaps somewhat less accurate because of the more complicated hyperfine structure and the greater Rydberg correction for $n\,^3D$ than for $n\,^3F$ states. We may consider the mean $T(2\,^3S_1) = 134044.24\,cm^{-1}$ as the final value based on triplets only. However, since the relative positions of the singlet and triplet levels is fixed (see Section D) we can obtain $T(2\,^3S_1)$ also from $T(2\,^1S)$ given above using the separation $2\,^3S_1$–$2\,^1S_0$ from Table IV and find $T(2\,^3S_1) = 134044.16\,cm^{-1}$. The agreement of the two values is very satisfactory and gives confidence in the determination of the series limits. The last value is probably somewhat more accurate since it is based on single lines. We give it twice the weight in forming the final average which is

$$T(2\,^3S_1) = 134044.19 \pm 0.10\,cm^{-1}.$$

This must be considered at present as the best available value. It may be compared with the best previous value, $134033\,cm^{-1}$, given by Moore (1949).

While we have not as yet remeasured the far ultraviolet resonance lines we may use the wave numbers given by Edlén as quoted by Robinson (1937) for $1\,^1S$–$2\,^1P$, $3\,^1P$, $4\,^1P$ together with our new term values for $2\,^1P$, $3\,^1P$, and $4\,^1P$ in order to obtain a revised value for the ionization potential of Li^+. We find

$$\text{I.P. } (Li^+) = 501806._5 + 108270.8_1 = 610077._3\,cm^{-1},$$

$$\text{or I.P. } (Li^+) = 561753._4 + 48326.5_8 = 610080._0\,cm^{-1},$$

$$\text{or I.P. } (Li^+) = 582835._5 + 27249.2_9 = 610084._8\,cm^{-1} \text{ respectively.}$$

Using weights corresponding to Robinson's quoted error limits we obtain as final mean

$$\text{I.P.}_{exp}(Li^+) = 610079._4\,cm^{-1}.$$

It appears that the major part of the previously quoted error ($\pm 25\,cm^{-1}$) was due to the uncertainty of the exact values of the term values of the excited states. The error quoted by Robinson for the wave number of the resonance lines was $\pm 5\,cm^{-1}$ and this is presumably the limit of error of the ionization potential. The excellent agreement of the three independent values suggests that the probable error of the mean may be less, perhaps $\pm 3\,cm^{-1}$. Indeed in a private communication received after the preceding discussion was written, Edlén has pointed out that the error estimate of the wavelengths of the resonance lines was a conservative one. He also emphasized that in selecting standards for these measurements no use was made of hydrogen-like spectra calculated without knowledge of the Lamb shift, an error that does affect several other wavelength determinations in the region below 200 Å.

F. CONCLUSION

Hart and Herzberg (1957) calculated the energy of the ground state of Li^+ by means of a 20-parameter Hylleraas function and obtained (including a relativistic correction of 19.6 cm^{-1})

$$I.P._{rel}^{20}(^7Li^+) = 610085.6\ cm^{-1}.$$

These calculations were much improved by Pekeris (1958), who went up to 203 parameters with certain approximations and obtained

$$I.P._{rel}^{203}(^7Li^+) = 610087.44\ cm^{-1}$$

a value that is estimated to be accurate to $\pm 0.05\ cm^{-1}$. Subtracting this Dirac value from the observed value, one obtains as an experimental value for the Lamb shift

$$\Delta E_{Lamb}^{(203)}(^7Li^+) = -8._0\ cm^{-1}$$

with an error ± 5 or possibly $\pm 3\ cm^{-1}$. The predicted Lamb shift is $-7.8\ cm^{-1}$ (Pekeris 1958). The agreement is better than might have been expected on the basis of the error in the observed value.

For the $2\ ^3S$ state until now only a 27-parameter non-relativistic energy value has been calculated (Hart and Herzberg 1959)

$$T_{n.r.}^{27}(2\ ^3S, Li^+) = 134029.7\ cm^{-1}.$$

The difference of this value and the observed value, 14.5 cm^{-1}, is the sum of the relativistic correction and the Lamb shift. Once the relativistic correction has been calculated an observed Lamb shift with an accuracy $\pm 0.10\ cm^{-1}$ will be obtainable.

While in the present paper the excited energy levels have been established with adequate accuracy, a definitive determination of the ground state remains to be done. For this purpose the far ultraviolet resonance lines will have to be measured with an accuracy of better than 0.001 Å. As far as the present results go, they confirm the predictions of quantum electrodynamics, as did those on He.

NOTE ADDED IN PROOF

Professor B. Edlén has called our attention to the fact that the Rydberg correction of the $2\ ^1S$ state does not fit very well with the Rydberg corrections of the higher S states and that Werner (1927) in his full paper (not accessible to us at the time of writing our paper) has given a different value for the $2\ ^1S$ state, which is based on a line at 1420.7 Å, assigned by him as $2\ ^1S$–$3\ ^1P$. This line was not mentioned in Werner's Notes in *Nature* and was not observed in our work. If Werner's identification of this line is correct, it would imply that the infrared line 8517.4 Å found independently by Series and Willis (1958) and ourselves is incorrectly assigned as $2\ ^1S$–$2\ ^1P$ and that the correct line must be at 9572 Å. The search for the infrared line, both in Series and Willis' work and our own, was based on an extrapolated value for the $2\ ^1S$ state given by Werner (1926) and quoted by Moore (1949). We did not search

near 9572 Å because of the much lower sensitivity of the infrared plates in that region. While the Rydberg correction of the $2\,^1S$ state based on our identification is indeed somewhat large, the Rydberg correction based on Werner's identification of the 1420 Å line is somewhat small. At the present moment it is difficult to decide which value is the correct one. It is even possible that neither one of the two values is correct. None of the remaining conclusions of our paper is affected by this uncertainty in the $2\,^1S$ state.

ACKNOWLEDGMENTS

We are greatly indebted to Mr. J. Shoosmith for his efficient help in obtaining the vacuum ultraviolet spectra and to Miss L. L. Howe for her meticulous work in the extensive computations. Sincere thanks are also due to Dr. A. E. Douglas for many helpful suggestions and critical remarks. We are grateful to Professors B. Edlén and A. G. Shenstone for a critical reading of the manuscript and a number of helpful suggestions.

REFERENCES

BETHE, H. A. and SALPETER, E. E. 1957. Quantum mechanics of one- and two-electron atoms (Academic Press, Inc., New York).
BOCKASTEN, K. 1955. Arkiv Fysik, **9**, 457.
DATTA, S. and BOSE, P. C. 1935. Z. Physik, **97**, 321.
DOUGLAS, A. E. and HERZBERG, G. 1957. J. Opt. Soc. Am. **47**, 625.
EDLÉN, B. and RISBERG, P. 1956. Arkiv Fysik, **10**, 553.
FRANCE, R. W. 1930. Proc. Roy. Soc. **129**, 354.
FRED, M., TOMKINS, F. S., BRODY, J. K., and HAMERMESH, M. 1951. Phys. Rev. **82**, 406.
GOUDSMIT, S. and BACHER, R. F. 1933. Phys. Rev. **43**, 894.
GRANATH, L. P. 1932. Phys. Rev. **42**, 44.
GÜTTINGER, P. and PAULI, W. 1931. Z. Physik, **67**, 743.
HART, J. F. and HERZBERG, G. 1957. Phys. Rev. **106**, 79.
——— 1959. Unpublished.
HERZBERG, G. 1958. Proc. Roy. Soc. **248**, 309.
JOHANSSON, I. 1959. Arkiv. Fysik, **15**, 169.
KABIR, P. K. and SALPETER, E. E. 1957. Phys. Rev. **108**, 1256.
KOPFERMANN, H. 1958. Nuclear moments (Academic Press, Inc., New York).
MOORE, C. E. 1949. Atomic energy levels, Vol. I (N.B.S. Washington, Circular 467).
PASCHEN, F. 1923. Ann. Physik, **71**, 537.
——— 1932. Sitzber. Akad. Berlin Phys.-Math. Kl. 502.
——— 1935. Sitzber. Akad. Berlin, 430.
PASCHEN, F. and CAMPBELL, J. S. 1938. Ann. Physik, **31**, 29.
PASCHEN, F. and RITSCHL, R. 1933. Ann. Physik, **18**, 867.
PEKERIS, C. L. 1958. Phys. Rev. **112**, 1649.
——— 1959. In press.
RISBERG, P. 1955. Arkiv Fysik, **9**, 483.
——— 1956. Arkiv Fysik, **10**, 583.
ROBINSON, H. A. 1937. Phys. Rev. **51**, 14.
SAWYER, R. A. and PASCHEN, F. 1927. Ann. Physik, **84**, 1.
SCHÜLER, H. 1924. Naturwiss. **12**, 579.
——— 1925. Ann. Physik, **76**, 292.
——— 1926. Z. Physik, **37**, 568.
——— 1927. Z. Physik, **42**, 487.
——— 1930. Z. Physik, **66**, 431.
SERIES, G. W. and WILLIS, K. 1958. Proc. Phys. Soc. **71**, 274.
SUCHER, J. 1958. Phys. Rev. **109**, 1010.
SUWA, S. 1950. Phys. Rev. **78**, 810.
——— 1951. J. Phys. Soc. Japan, **6**, 231.
TOMKINS, F. S. and FRED, M. 1951. J. Opt. Soc. Am. **41**, 641.
WERNER, S. 1925. Nature (London), **115**, 191; **116**, 574.
——— 1926. Nature (London), **118**, 154.
——— 1927. Ph.D. Thesis, Copenhagen.

PRINCIPE D'UNE ANALYSE DES COURBES EXPÉRIMENTALES COMPOSÉES: APPLICATION À LA MESURE DE LA DENSITÉ DES GRAINS DANS LES TRACES FORTEMENT IONISÉES DANS LES ÉMULSIONS NUCLÉAIRES[1]

W. Paskievici[2]

RÉSUMÉ

On expose le principe d'une méthode mathématique, utilisant les transformées de Fourier, qui rend possible la décomposition, dans ses éléments constitutifs, d'une courbe expérimentale complexe, lorsque celle-ci provient de la superposition de plusieurs courbes ayant toutes la même forme analytique, mais des amplitudes différentes. On considère, en particulier, la possibilité d'utiliser cette analyse pour déterminer, à partir des courbes photométriques, la densité des grains des traces laissées par les particules fortement ionisantes dans les émulsions ionographiques. On montre dans un exemple simple, idéalisé, l'utilisation et la précision de la méthode.

INTRODUCTION

Dans le présent compte rendu on expose le principe d'une méthode d'analyse des courbes expérimentales complexes lorsque celles-ci présentent des caractères mathématiques simples. Cette méthode peut être utilisée, en général, dans tout problème physique, lorsqu'il s'agit de décomposer dans ses éléments constitutifs, une courbe qui provient de la superposition de plusieurs courbes ayant toutes la même forme, mais des hauteurs différentes.

Le premier exemple d'application de cette méthode est celui de l'analyse des courbes photométriques représentant l'opacité des trajectoires des particules ionisantes dans une émulsion ionographique.

Cette analyse a pour but de déterminer la densité des grains développés et leur opacité ou degré de noircissement dans la partie finale de la trajectoire lorsque les mesures visuelles deviennent imprécises ou même impossibles.

Ce mode d'analyse a été conçu pour s'appliquer au modèle "optique" des grains, dans l'hypothèse "d'invariance" d'une "famille simplement définie" (Ahmad et Demers 1959). Dans cette hypothèse, les courbes photométriques des grains ont la même forme analytique à une constante multiplicative près; de plus, lorsque les grains sont suffisamment rapprochés, les courbes photométriques montrent une région de superposition sans interférence, avec simple addition des densités optiques. Nous renvoyons à l'article de ces auteurs pour la justification de ces hypothèses, qui paraissent rendre compte de façon très convenable des observations photométriques: absence d'interférences perceptibles et possibilité d'appliquer le modèle même pour des grains écartés de l'axe.

[1]Manuscrit reçu le 10 juin 1959.
Contribution du Département de Physique de la Faculté des Sciences de l'Université de Montréal, Montréal, Qué.
[2]Boursier postdoctoral du Conseil National des Recherches. Adresse actuelle: École Polytechnique, Université de Montréal, Montréal, Qué.

FORMULATION DE LA MÉTHODE

Il s'agit d'analyser une courbe qui est mathématiquement exprimée par la relation:

$$F(x) = \sum_i A_{\xi_i} S(x-\xi_i). \tag{1}$$

Dans le cas physique actuellement sous considération, $F(x)$ représente la courbe photométrique, ou encore l'opacité ou degré de noircissement de la trace selon son axe x, $S(x-\xi_i)$ sont des courbes de la même famille dont l'allure représente l'opacité des grains individuels le long de l'axe de la trajectoire. Ces courbes sont symétriques par rapport aux points $x = \xi_i$ qui sont les centres des grains i, leur forme analytique est laissée arbitraire, mais supposée connue dans les applications; une bonne approximation semble être

$$S(x-\xi_i) = \frac{\sin^n(x-\xi_i)}{(x-\xi_i)^n} \tag{2}$$

avec $n \geqslant 2$. Définissons une fonction de distribution des centres des grains

$$D(\xi) = \sum_i A_{\xi_i} \delta(\xi-\xi_i) \tag{3}$$

ou, si on pose $A(\xi_i) = A_{\xi_i}$:

$$D(\xi) = \sum_i A(\xi)\delta(\xi-\xi_i). \tag{3'}$$

Alors, à cause des propriétés des fonctions δ de Dirac,

$$F(x) = \int_0^\infty D(\xi)S(x-\xi)d\xi. \tag{4}$$

La limite supérieure de l'intégrale est prise par excès sans introduire par cela aucune erreur.

En prenant la transformée cosinus de Fourier de chaque membre de l'égalité (4) et en tenant compte du théorème des convolutions des fonctions transformées de Fourier (Morse et Feshbach 1953), on obtient la relation simple:

$$f(t) = \sqrt{2\pi}\, d(t)\, s(t) \tag{5}$$

où

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty F(x) \cos xt\, dx, \\ d(t) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty D(\xi) \cos \xi t\, d\xi, \\ s(t) &= \sqrt{2/\pi} \int_0^\infty S(r) \cos rt\, dr \end{aligned} \tag{6}$$

où $r = x-\xi$. Par conséquent,

$$d(t) = \frac{1}{\sqrt{2\pi}} \frac{f(t)}{s(t)} = \frac{1}{\sqrt{2\pi}} g(t). \tag{7}$$

Si on effectue maintenant la transformée inverse, on a

$$D(\xi) = \frac{1}{2\pi} \int_0^\infty g(t) \cos \xi t \, dt. \tag{8}$$

On pourra donc obtenir la distribution des grains $D(\xi)$ à partir de la transformée cosinus du quotient des transformées cosinus des courbes photométriques d'une portion de la trace $F(x)$ et d'un grain individuel, $S(r)$.

$D(\xi)$ est, dans les hypothèses admises au départ, une somme de fonctions δ, représentables physiquement par une série de traits verticaux placés sur l'axe x où chacun représente le centre d'un grain, la hauteur de ces traits prise égale au coefficient A_{ξ_i} signifiant l'opacité des grains en question.

À la suite de l'opération mathématique, la fonction expérimentale continue $F(x)$ se trouve remplacée par une fonction discontinue $D(\xi)$. Si l'hypothèse "d'invariance" n'est qu'une approximation, ce qui est probable, la fonction $D(\xi)$ résultante sera certainement modifiée. Il est difficile d'imaginer d'avance ces modifications sans calculs numériques précis; on peut espérer que la courbe photométrique se trouvera alors remplacée, sinon par une série de traits infiniment étroits, du moins par une nouvelle courbe continue où la position et l'opacité des grains seront plus facile à reconnaître.

DISCUSSION

Dans le présent compte rendu, on exposera seulement quelques considérations sur la résolution mathématique du problème.

Il faudra d'abord déterminer la transformée $s(t)$ qui sera la même pour toutes les traces des plaques d'une même émulsion; en attribuant à $S(r)$ une forme analytique convenable, on pourra facilement calculer ou obtenir des tables, $s(t)$.

Le calcul de $f(t)$ est difficile car $F(x)$ ne peut pas être approchée par une fonction analytique. Généralement il faudra utiliser un calculateur électronique; dans certains cas, cependant, on pourrait songer à un développement en polynômes de Legendre:

$$F(x) = \sum_l \sqrt{\pi/2}\, a_l\, (-1)^l P_{2l}(x/x_0) \qquad |x| < x_0.$$

Une fois obtenus les coefficients a_l on calcule aisément la transformée

$$f(t) = \sum_l a_l j_{2l}(tx_0)$$

avec $j_l(x)$, les fonctions sphériques de Bessel habituelles.

Il est difficile de prévoir l'allure de $g(t)$ sans effectuer les calculs; on peut affirmer cependant qu'elle ne devrait pas être singulière car, en général, si $s(t)$ s'annule, $f(t)$ s'annule aussi.

Par conséquent, la seule difficulté théorique qui apparaît provient du fait que $D(\xi)$ est une fonction de Dirac ce qui rend problématique l'utilisation d'un calculateur électronique. Cependant le calcul numérique est encore possible car, grâce à l'intervalle d'intégration limité, on obtient une fonction

approchée de celle de Dirac et qui tend vers celle-ci lorsque l'intervalle d'intégration augmente indéfiniment.

Afin de vérifier si l'utilisation de la méthode à l'aide du calculateur était possible et de connaître la précision des résultats obtenus, on a commencé la programmation d'une série de calculs en partant des valeurs numériques d'une courbe construite synthétiquement à ce propos.

APPLICATION NUMÉRIQUE

La courbe synthétique (figure 1) a été composée de huit courbes individuelles ayant la forme

$$S(x-\xi_i) = \frac{\sin^2 a(x-\xi_i)}{a^2(x-\xi_i)^2} \qquad \text{pour } \left(\xi_i - \frac{\pi}{a}\right) \leqslant x \leqslant \left(\xi_i + \frac{\pi}{a}\right)$$

$$= 0 \qquad \text{hors de cet intervalle.}$$

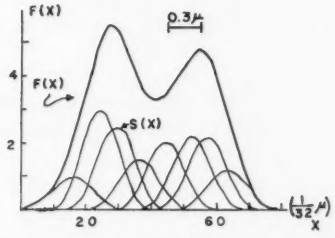


FIGURE 1.

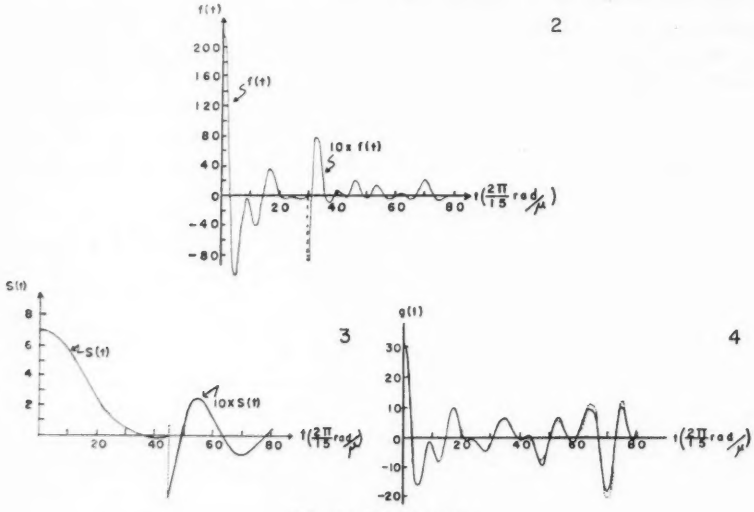
Le coefficient a dépend de l'échelle de longueur utilisée; dans notre cas, où l'unité de longueur est égale à $1/32\,\mu$, $a = 6.4\,\mu^{-1}$. Dans le tableau I on donne les valeurs de ξ_i, dans la même unité, et de A_{ξ_i}, en unités arbitraires; le choix de ces valeurs correspond du point de vue physique à une trace ionisante complètement saturée, lorsque le comptage des grains devient impossible.

TABLEAU I

i	ξ_i	A_{ξ_i}	i	ξ_i	A_{ξ_i}
1	16	1.0	5	45	2.0
2	25	3.0	6	53	2.2
3	30	2.5	7	58	2.2
4	37	1.5	8	64	1.2

La courbe $F(x)$ a été calculée par points, pour 80 valeurs, en utilisant des tables à quatres décimales exactes; on voit qu'elle présente uniquement deux maxima. Les calculs numériques ont été effectués à l'aide d'un calculateur arithmétique Librascope LGP-30 de Royal McBee Corporation; pour évaluer les intégrales, on a utilisé la méthode de Simpson.

Les figures 2, 3 et 4 représentent, respectivement, les transformées $f(t)$, $s(t)$ et leur quotient $g(t)$; dans la dernière figure on a tracé aussi, en pointillé,



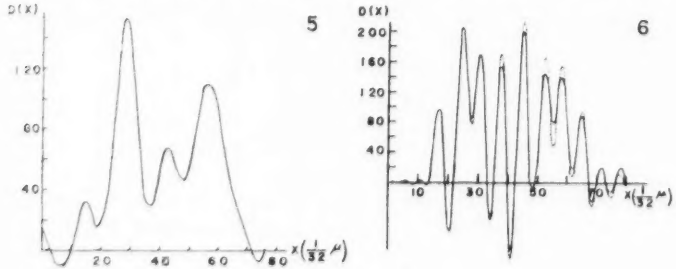
FIGURES 2,* 3* et 4.

la fonction $g(t)$ calculée exactement à partir de la fonction connue dans notre cas, $D(\xi)$ dont elle est la transformée. Plus explicitement

$$g_{\text{exacte}}(t) = \sum_{\xi} A_{\xi} \cos \xi t.$$

En comparant ces deux dernières courbes, on constate une excellente précision jusque vers $t = 40$, ce qui correspond à la première moitié de l'intervalle considérée, tandis qu'elle diminue légèrement dans la seconde moitié, comme il fallait s'y attendre d'ailleurs, à cause de l'erreur croissante sur l'évaluation de la transformée de Fourier pour t élevé.

Finalement, dans les figures 5 et 6 on montre la transformée de Fourier



FIGURES 5 et 6.

*Dans les figures 2 et 3, au lieu de $10 \times f(t)$ lire $100 \times f(t)$, au lieu de $10 \times S(t)$ lire $100 \times S(t)$.

de $g(t)$ définie jusqu'à t égal à 40 et à 80 respectivement. Dans cette dernière figure, dans laquelle on a indiqué aussi, en pointillé, la transformée de Fourier de $g(t)$ exacte, on observe aisément les huit pics principaux qui correspondent aux huit courbes initiales. L'analyse de la courbe $F(x)$ est ainsi réalisée.

Il est important de remarquer que la courbe de la figure 6, devrait être $D(\xi)$ mais qu'elle n'est qu'une approximation d'une somme de fonctions de Dirac δ; ceci est dû au fait que si l'intervalle d'intégration sur t est limité à une certaine valeur t_0, on obtient pour $D(\xi)$

$$D(\xi) = \sum_i \frac{\sin (\xi - \xi_i) t_0}{(\xi - \xi_i)}.$$

Si t_0 est suffisamment grand, comme dans notre cas où $t_0 = 80$, la séparation entre les courbes est suffisante pour déterminer avec précision le nombre de maxima et leur emplacement; les valeurs des maxima qui sont proportionnelles aux coefficients d'opacité A_{ξ_i} ne sont pas déterminées avec la même précision parce que les courbes individuelles présentent encore des domaines d'interférence.

CONCLUSION

La méthode développée ci-dessus devrait permettre, dans l'esprit de l'auteur, de décomposer une courbe, qui est une superposition de plusieurs courbes individuelles, dans ces éléments constitutifs. Les résultats obtenus à partir d'une courbe synthétique, mais se rapprochant autant que possible de la courbe réelle, montrent que la décomposition est mathématiquement possible.

Une application immédiate est en vue: celle d'obtenir la position et par suite la densité des grains et leur opacité, le long des traces ionographiques de forte densité, lorsque le comptage habituel devient impossible.

L'analyse des profils expérimentaux présente des difficultés particulières dues essentiellement au caractère approché de l'hypothèse de base qui considère tous les grains de même diamètre et parfaitement alignés selon un axe, et aux erreurs expérimentales. L'influence de ces facteurs dans le procédé d'analyse présentement développé sera étudiée dans une publication ultérieure.

La connaissance de ces valeurs contribuera à l'étude de la structure des traces, de la perte d'énergie par ionisation, à la discrimination plus rapide et peut-être plus précise des particules; elle pourra vérifier, en particulier, certaines hypothèses avancées sur l'existence des grains "gris" associés au phénomène de contagion dans l'émulsion ionographique.

REMERCIEMENTS

Ce travail fait partie du programme de recherches que j'effectue à l'Université de Montréal, grâce à une bourse postdoctorale du Conseil National de Recherches.

Je remercie Monsieur le Professeur P. Demers pour m'avoir proposé le sujet actuellement sous étude, qui m'a aidé à comprendre certains aspects de la structure des traces ionographiques et des problèmes d'identification des particules nucléaires.

Je remercie aussi Monsieur I. Ahmad pour les nombreuses discussions que j'ai eues avec lui sur ce sujet; son travail de thèse m'a beaucoup aidé.

Je remercie également Monsieur B. Banville qui s'est chargé de la programmation des calculs; sans son aide je n'aurais pu obtenir les résultats numériques actuels.

NOTE AJOUTÉE SUR ÉPREUVE

La méthode a maintenant été utilisée avec succès pour séparer deux grains très rapprochés dans un "paquet". On s'est basé sur les courbes photométriques d'un grain isolé et d'un "paquet", obtenues dans une émulsion à grain fin préparée au laboratoire. Je remercie M. Ivan Batthyany, pour son aide dans ses dernières mesures.

RÉFÉRENCES

AHMAD, I. et DEMERS, P. 1959. Phot. Corp. II (Les Presses Universitaires de Montréal).
AHMAD, I. 1958. Structure et identification des trajectoires dans les émulsions ionographiques à grain fin. Thèse de Doctorat, Université de Montréal, Montréal, Québec.
MORSE, P. et FESHBACH, H. 1953. Methods of theoretical physics (McGraw-Hill Book Co., Inc.), p. 465.

NOTES

SOME OBSERVATIONS ON NUCLEAR ALIGNMENT IN COBALT METAL

J. M. Daniels and M. A. R. LeBlanc

The possibility of obtaining nuclear alignment using the interactions in a ferromagnetic single crystal was demonstrated by Grace *et al.* (1955, 1957), who showed that when a cobalt metal single crystal containing a little Co 60 is cooled to about 0.05° K, a spontaneous nuclear alignment occurs parallel to the hexagonal axis of the cobalt crystal. This alignment causes the γ-rays from the Co 60 to be emitted anisotropically with respect to the hexagonal axis of the cobalt crystal. The mechanism which produces this alignment can be described, qualitatively, as follows. Cobalt metal crystallizes in the hexagonal system; it is also ferromagnetic and the direction of spontaneous magnetization of the domains is parallel to the hexagonal axis below 250° C. Thus, at low temperatures, all the unpaired electron spins responsible for the ferromagnetism are aligned parallel to the hexagonal axis, and each spin produces a magnetic field of several hundred thousand gauss, parallel to this axis, at the site of the nucleus to which it is attached. At temperatures below 1° K, the cobalt nuclei are aligned parallel to the hexagonal axis, because this orientation describes the state of lowest energy, which is preferentially populated. We have aligned Co 60 in a cobalt metal single crystal by this method, and have made measurements of the anisotropy of γ-rays emitted by the aligned Co 60 nuclei. Our results differ significantly from those of other workers, and an account of them is presented below, with a possible explanation.

A piece of cobalt metal single crystal, $0.13 \times 0.13 \times 0.30$ cm in size, was given to us by Dr. P. Myers, formerly of the Department of Mining and Metallurgy, University of British Columbia. It was irradiated to an activity of about 3 μc in the neutron flux of N.R.X., Chalk River. It was cooled to temperatures of the order of 0.04° K, by being placed in contact with pieces of potassium chromium alum, cooled to this temperature by adiabatic demagnetization. The temperature of this potassium chromium alum was measured by the ballistic method. The intensity of emission of γ-rays was measured by scintillation counters placed along, and perpendicular to, the hexagonal axis of the cobalt crystal. This apparatus has been described in greater detail by Daniels *et al.* (1958). Various methods were used to ensure good heat contact between the cobalt metal and the chromium potassium alum. These were:

(i) The crystal was placed between two cylinders, $\frac{3}{8}$ in. diameter $\times$ 1 in. long, of compressed chrome alum, and the gap between was filled with glycerine, with a solution of chrome alum in 50% water, or with a sludge of Apiezon oil B and chrome alum powder.

(ii) The crystal was pushed into a hole cut in a copper strip 3 cm×1 cm and the whole assembly was copper-plated, bonding the crystal to the copper fin. This was sandwiched between four half cylinders of chrome alum, and lubricated with a sludge of chrome alum and glycerine.

(iii) About 200 No. 36 B&S copper wires providing a surface area of about 150 cm² were soldered to the fin. Chrome alum powder moistened with Apiezon oil B was carefully placed between the wires, and the whole assembly compressed into a solid in a pill press.

(iv) A thin copper foil of about 200 cm² surface was soldered to the copper fin, and folded in a concertina shape. Chrome alum powder and Apiezon oil B was placed between the folds and pushed down hand-tight. The assembly was held in a lucite container.

The results obtained from all these methods were self-consistent during the time that the anisotropies were decreasing and the chrome alum was warming. Hence we are satisfied that there was adequate heat contact between the chrome alum and the cobalt crystal in all our experiments.

The anisotropy of emission of γ-rays, ϵ, is defined by

$$\epsilon = \left\{ I\left(\frac{\pi}{2}\right) - I(0) \right\} \Big/ I\left(\frac{\pi}{2}\right)$$

where $I(0)$ and $I(\pi/2)$ are respectively the counting rates along and perpendicular to the hexagonal axis of the cobalt crystal. The results of a typical experiment are given in the figure. If we assume that after about an hour the cobalt crystal has come into thermal equilibrium with the chrome alum,

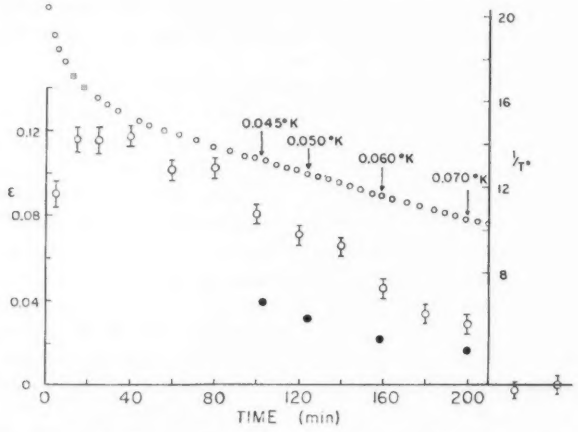


FIG. 1. The measured anisotropy of γ-ray emission from cobalt 60 in cobalt metal single crystal, and the temperature of the specimen, both as a function of time. This is for our cobalt crystal, after heat treatment, cooled by 150 cm² surface of No. 36 B&S copper wires as described in the text. The solid points represent the values obtained at Oxford for the temperatures indicated.

we can correlate ϵ with the absolute temperature T. At high temperatures (above about 0.05° K in this case) $\epsilon = a/T^2$ where a is a constant which describes, among other things, the hyperfine coupling of the Co 60 nucleus with the electron spins.

We observed $a = 3\times10^{-4}$. This is considerably larger than the Oxford values, $a = 0.79\times10^{-4}$ (Kurti 1959), and rather smaller than the value of $a \simeq 5\times10^{-4}$ reported by Khutsishvili (1955). Now the hyperfine coupling is responsible not only for the anisotropy of γ-rays, but also for the specific heat below 1° K. This has been measured by Heer and Erickson (1957), Heer (1958), and Arp *et al.* (1957). Heer and Erickson give for the specific heat C, $CT^2/R = 4\times10^{-4}$, and Arp *et al.* give $CT^2/R = 4.7\times10^{-4}$. If we assume a hyperfine coupling of the form $AS_zI_z+B(S_xI_x+S_yI_y)$ we have

$$\frac{CT^2}{R} = (A_{59}^2+2B_{59}^2)I(I+1)/12 \qquad \text{and} \qquad \epsilon = 39(A_{60}^2-B_{60}^2)/28k^2T^2$$

where the subscripts 59 and 60 refer to the isotopes Co 59 and Co 60 respectively. The B term is negligible, hence we can calculate what value of a we should expect from the measured values of specific heat. The two values quoted give, respectively, $a = 1.4\times10^{-4}$ and 2.17×10^{-4}.

The crystal was heated to a red heat for a few minutes in air and the experiments repeated. This time the anisotropies were much smaller, $a = 1.5\times10^{-4}$. We suggest the following explanation for the discrepancy between all the values quoted for a. Cobalt has a phase transformation at 417° C, the low-temperature stable phase being hexagonal, and the high-temperature phase cubic. As a result, it is almost impossible to prepare single crystals of hexagonal cobalt, small cubic regions being present as stacking faults. A proportion of 10% of cubic phase is quite common (Houska and Averbach 1958), and proportions of more than 50% have been observed in supposedly hexagonal cobalt (Edwards and Lipson 1942). If we assume, as seems reasonable, that our crystal contained more of the hexagonal phase before heat treatment than after, then we must assume that the hyperfine coupling in the cubic phase is much less than in the hexagonal phase. This would explain why different experimenters have obtained different results for a and also why the different measurements of specific heat of cobalt metal differ by amounts which are apparently much greater than the experimental error. It would appear that the crystal used by Khutsishvili contained more of the hexagonal phase than did ours, and that ours contained more of the hexagonal phase than did the crystal used at Oxford. It would be of interest to measure both the specific heat and the anisotropy of γ-rays from cobalt metal for which the proportions of cubic and hexagonal phases are known. The proportions can be measured by X-ray analysis, however, we have not the necessary equipment for this at our disposal.

The support for this work of the National Research Council and of Atomic Energy of Canada Limited is gratefully acknowledged.

ARP, V., KURTI, N., and PETERSON, R. 1957. Bull. Am. Phys. Soc. **2**, 388.
DANIELS, J. M., LAMARCHE, J. L. G., and LEBLANC, M. A. R. 1958. Can. J. Phys. **36**, 997.
EDWARDS, O. S. and LIPSON, H. 1942. Proc. Roy. Soc. (London), A, **180**, 268.
GRACE, M. A., JOHNSON, C. E., KURTI, N., SCURLOCK, R. F., and TAYLOR, R. T. 1955. Communications sur la Conférence de Physique des Basses Températures, Paris (Institut International du Froid), p. 263.
——— 1957. Bull. Am. Phys. Soc. **2** (II), 136.
HEER, C.V. 1958. Program K. O. Conference, 36.
HEER, C.V. and ERICKSON, R.A. 1957. Phys. Rev. **108**, 896.
KHUTSISHVILI, G. R. 1955. J. Exptl. Theoret. Phys. (U.S.S.R.), **29**, 894.
KURTI, N. 1959. J. phys. radium, **20**, 141.
HOUSKA, C. R. and AVERBACH, B. L. 1958. Acta Cryst. **11**, 139.

RECEIVED AUGUST 14, 1959.
DEPARTMENT OF PHYSICS,
UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, BRITISH COLUMBIA.

NUCLEAR SPIN RELAXATION IN LOW-TEMPERATURE LIQUIDS

IRWIN OPPENHEIM* AND MYER BLOOM

The nuclear spin lattice relaxation time T_1 may be expressed in terms of the correlation functions of those interactions which enable energy to be exchanged between the nuclear spin system and the other degrees of freedom of the material in which the spins are located (Bloembergen, Purcell, and Pound 1948). Measurements of T_1 are therefore capable of giving information related to the statistical mechanical properties of matter. In this note we discuss the interpretation of T_1 measurements in liquids, and particularly in low-temperature liquids, where the potentialities of the nuclear magnetic resonance technique have not yet been realized.

There are two main types of relaxation mechanisms to be considered.

I. In the cases of atomic liquids, such as helium, xenon, etc., the spin relaxation is caused by the dipole–dipole interaction between neighboring nuclear spins and the well-known result for T_1 is expressed in terms of the correlation functions of $Y_{2m}(\theta, \phi)/R^3$ where R, θ, ϕ are the polar co-ordinates of the vector joining a pair of atoms (Bloembergen, Purcell, and Pound 1948).

II. In the cases of molecular liquids, such as H_2, CH_4, H_2O, etc., in which the molecules have rotational angular momentum $\mathbf{J}$ an additional mechanism is present due to interactions within the molecules which couple the nuclear spins to $\mathbf{J}$. Here the problem is to calculate the correlation function $\langle \mathbf{J}(t)\mathbf{J}(t+\tau)\rangle$. The simplest case is H_2 at low temperatures where all the ortho-H_2 molecules are in the $J = 1$ state and $\langle \mathbf{J}(t)\mathbf{J}(t+\tau)\rangle$ may be written in terms of the mean time spent by a molecule in an m_J state, which in turn depends on the correlation functions of the anisotropic part of the intermolecular potential (Bloembergen, Purcell, and Pound 1948; Moriya and Motizuki 1957; Moriya 1957).

*National Bureau of Standards, Washington, D.C.

As will be shown in a future publication, for many systems a plausible approximation for the correlation function $\langle\phi(t)\phi(t+\tau)\rangle$, in which $\phi(t)$ is a function only of the vector $\mathbf{R}(t)$ joining the pairs and not of their momenta, is

$$(1) \qquad \langle\phi(t)\phi(t+\tau)\rangle \cong \iint d^3\mathbf{R}d^3\mathbf{R}'[g(R)]^{1/2}[g(R')]^{1/2}P(\mathbf{R}, \mathbf{R}', \tau)\phi(\mathbf{R})\phi(\mathbf{R}')$$

where $P(\mathbf{R}, \mathbf{R}', \tau)$ is related to the probability per unit volume that two particles separated by $\mathbf{R}$ at a time $t = 0$ are separated by $\mathbf{R}'$ at a time τ. Note that the form in which the radial distribution function $g(R)$ appears in equation (1) is slightly different from the usual form in which it appears in the correlation function (Torrey 1958). For a liquid we assume the particles obey the diffusion equation (Chandrasekhar 1943) in which case

$$(2) \qquad P(\mathbf{R}, \mathbf{R}', \tau) = (1/8\pi D\tau)^{3/2} \exp[-(\mathbf{R}-\mathbf{R}')^2/8D\tau]$$

where D is the diffusion constant.

$g(R)$ has been evaluated using the superposition approximation for liquids in which molecules interact via a Lennard–Jones potential (Kirkwood, Lewinson, and Alder 1952). In applying these results to low-temperature systems, however, one should not use the true temperature of the liquids T but the "effective temperature" T_{eff} defined by $(3/2)k_B T_{\text{eff}} = \langle p^2/2m\rangle$ (Mazo and Kirkwood 1955; Oppenheim 1954). This is an important point since in low-temperature liquids, such as He^3 and H_2, T_{eff} increases as T decreases. For H_2 $T_{\text{eff}}(14°\text{ K})/T_{\text{eff}}(20°\text{ K}) = 1.18$, while for He^3 $T_{\text{eff}}(1°\text{ K})/T_{\text{eff}}(3°\text{ K}) = 1.42$.

The final results are:

Case I

$$(3) \qquad \frac{1}{T_1} = \frac{4\pi\gamma^4\hbar^2 I(I+1)N}{D}\int_0^\infty \frac{dk}{k}\left[\int_0^\infty \frac{[g(R)]^{1/2}J_{5/2}(kR)}{R^{3/2}}\,dR\right]^2;$$

Case II

$$(4) \qquad T_1 \propto \frac{1}{\tau_1} \propto \frac{N}{D}\int_0^\infty \frac{dk}{k}\sum_{l,m}\left[\int_0^\infty [g(R)]^{1/2}R^{3/2}a_{lm}(R)J_{l+\frac{1}{2}}(kR)dR\right]^2,$$

where N = number of molecules per cubic centimeter and γ is the nuclear gyromagnetic ratio. We have used the fact that the correlation times for self-diffusion are short compared with the Larmor period. The anisotropic interactions between molecules have been expanded as

$$\phi(\mathbf{R}) = \sum_{l,m} a_{lm}(R)\,Y_{lm}(\theta, \phi);$$

$J_{l+\frac{1}{2}}(kR)$ denotes a Bessel function of order $l+\frac{1}{2}$. τ_1 is the mean lifetime of the molecule in an m_J state.

The best technique for measuring T_1 is the pulse method of Hahn (1950). This method has the additional advantage that it enables a direct and reliable measurement of D and hence a direct comparison of theory with experiment is possible. Unfortunately, however, although T_1 has been measured for H_2 liquid (Bloom 1957) no measurements of D were made.

In Fig. 1, we have plotted the temperature dependence of the diffusion constant of H_2 liquid required to explain the observed temperature dependence of T_1 (Bloom 1957) according to equation (4). In this plot we have neglected the temperature dependence introduced by the integral; the integral varies

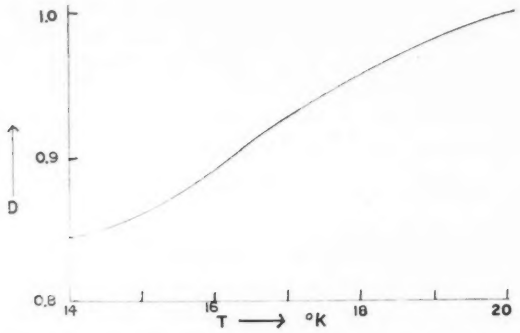


FIG. 1. Plot of predicted values of diffusion constant for H_2 liquid as a function of temperature using measured values of N and T_1. The diffusion constant is plotted in arbitrary units normalized to the value of 20° K.

by less than 5% between 20° K and 14° K for H_2 liquid in equilibrium with its vapor. The reason for this small effect is that as the temperature drops, T_{eff} and the density both increase and their effects on $g(R)$ tend to cancel each other (Kirkwood, Lewinson, and Alder 1952). This term will, however, be very important in calculating the strong dependence of T_1 in H_2 on ortho-H_2 concentration, the small difference in the temperature dependence of samples of H_2 having different ortho-H_2 concentrations (Bloom 1957) and the effect of pressure on T_1. Results of calculations bearing on these experiments will be presented in a subsequent publication. The reasonable temperature dependence of D required to explain the experimental behavior of T_1 according to this theory is to be contrasted with predictions of the Bloembergen–Purcell–Pound (BPP) theory for liquids which gives $T_1 \propto T/\eta$ where η is the viscosity. This expression predicts a decreasing T_1 with decreasing temperature instead of the observed increasing T_1. A similar temperature dependence to ours was obtained by Moriya (1957) using a different approach.

It is not at all surprising that the BPP theory breaks down completely at low temperatures for case II type relaxation. Their model of a sphere moving in a viscous medium used in calculating correlation functions for molecular reorientation may be a fair approximation only for high values of J when the reorientational motion may be described as a random walk process. For a liquid such as H_2, however, where only the $J = 1$ state is excited, the BPP model breaks down, as a single transition from one m_J state to another causes appreciable reorientation. Actually, all we have done in case II is to extend the BPP treatment of H_2 gas to the liquid case.

The results for He^3 liquid in equilibrium with its vapor are almost the same as the BPP theory since, again, the integral in equation (3) is almost temperature independent. We would therefore expect T_1 to decrease with decreasing temperature, according to measurements of D for He^3 (Garwin and Reich 1959). It is expected, however, that the Fermi–Dirac statistics of the He^3 atoms should become important and should cause T_1 to increase strongly at low temperatures upon the onset of degeneracy. We postpone a discussion of this difficult point since we are here only concerned with dynamical effects.

Finally, it may be remarked that experiments which would provide a strong test of this theory are the measurement of T_1 as a function of temperature and pressure in simple liquids such as xenon and neon, where effects of statistics are small. For these cases equation (3) should hold, and all the parameters appearing there are already known.

BLOEMBERGEN, N., PURCELL, E. M., and POUND, R. V. 1948. Phys. Rev. **73**, 679.
BLOOM, M. 1957. Physica, **23**, 378.
CHANDRASEKHAR, S. 1943. Revs. Modern Phys. **15**, 1.
GARWIN, R. L. and REICH, H. A. 1959. Phys. Rev. (in press).
HAHN, E. L. 1950. Phys. Rev. **80**, 980.
KIRKWOOD, J. G., LEWINSON, V. A., and ALDER, B. J. 1952. J. Chem. Phys. **20**, 929.
MAZO, R. M. and KIRKWOOD, J. G. 1955. Proc. Natl. Acad. Sci. **41**, 204.
MORIYA, T. 1957. Progr. Theoret. Phys. **18**, 183.
MORIYA, T. and MOTIZUKI, K. Progr. Theoret. Phys. **18**, 567.
OPPENHEIM, I. 1954. National Bureau of Standards Report.
TORREY, H. C. 1958. Nuovo cimento (Suppl. II), No. 1, 95.

RECEIVED JUNE 24, 1959.
DEPARTMENT OF PHYSICS,
UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, B.C.

LETTERS TO THE EDITOR

Under this heading brief reports of important discoveries in physics may be published. These reports should not exceed 600 words and, for any issue, should be submitted not later than six weeks previous to the first day of the month of issue. No proof will be sent to the authors.

Frequency Measurement of Standard Frequency Transmissions Against Caesium-Beam Resonator Standard*

Since July 1958 a caesium-beam frequency resonator has been in operation at the National Research Council laboratories (Kalra, Bailey, and Daams 1958, 1959). The local Essen ring quartz crystal oscillators (Kalra, Pattenson, and Thomson 1959) are calibrated periodically against this resonator standard. Since the day-to-day stability of these oscillators is of the order of $2{:}10^{10}$, these oscillators do not need a continuous monitoring against the caesium-beam frequency standard. An equation of the type $f = f_0(1+at)$ is computed each month for each oscillator, where f is the frequency on any day, a is a constant, and t is time in days.

Kalra, Pattenson, and Thomson (1959) have briefly noted the measurement of the carrier frequency of some of the very low frequency stations with precision-controlled carrier frequencies by a phase comparison method. These measurements are made against one of the

TABLE I

Frequency measurement of standard frequency transmissions

August 1959	MSF, 60 kc/s	GBR, 16 kc/s		KK2XEI, 60 kc/s
		6-hour average*	24-hour average	
1	−168	−167		N.M.
2	−167	−173		N.M.
3	−162	N.M.		−86
4	−158	−169		−83
5	−161	−163		−87
6	−160	−157		−87
7	−158	−161		−90
8	−161	−167		N.M.
9	−162	−164		N.M.
10	−155	−157		−86
11	−185	−170		−85
12	N.M.	−184		−85
13	−182	−185	−179	−79
14	N.M.	−176	−176	−83
15	−172	−176	−177	N.M.
16	−181	−176	−179	N.M.
17	−177	−169	−169	−81
18	−173	−172	−172	−81
19	−175	−175	−171	−81
20	−176	−174	−176	−81
21	−177	−177	−178	−81
22	−172	−178	−179	N.M.
23	−176	−178	−176	N.M.
24	−181	−178	−177	−82
25	−182	−179	−178	−78
26	−180	−175	N.M.	−79
27	−182	−176	N.M.	−79
28	−170	−178	−174	−80
29	−166	−173	−169	N.M.
30	−172	−170	−170	N.M.
31	−172	−166	−170	−76
Midmonthly mean	−171	−171	−176	−82
Midmonthly mean of WWV	−103			

NOTE: Measurements are made at Ottawa, Canada, using N.R.C. caesium-beam frequency resonator as reference standard (with an assumed frequency of 9 192 631 770 c.p.s.). Frequency deviations from nominal are quoted in parts per 10^{10}. A negative sign indicates that the frequency is below nominal.

N.M.: No measurement.

*Time of observations: 0100 to 0400 and 1600 to 1300 U.T.

*Issued as N.R.C. No. 5398.

Essen ring oscillators. By means of the computed equation above, these measurements are later referred to the caesium-beam frequency resonator standard. The time signals from some of the standard frequency or time stations are also monitored against these oscillators working as clocks. The time signals and the carrier frequencies of these stations are usually controlled by the same crystal oscillator. The carrier frequency of these stations is thus computed with an averaging period of 1 calendar month. For both techniques the over-all accuracy of the frequency averaged around midmonth is estimated to be $\pm 5{:}10^{10}$. Full details of the techniques will be published later.

The accompanying table shows the frequency of MSF (Rugby, England) 60 kc/s, GBR (Rugby, England) 16 kc/s, KK2XEI (Boulder, Colorado, U.S.A.) 60 kc/s, as measured by the phase technique; and the midmonthly mean frequency of WWV (Beltsville, Maryland, U.S.A.) 5 and 10 Mc/s, as measured against the N.R.C. caesium-beam resonator standard. Values are quoted in fractional parts in 10^{10}, i.e. $\Delta f/f \times 10^{10}$ where Δf is the departure from nominal and f the nominal frequency. The reference frequency scale is defined such that the N.R.C. caesium-beam resonator has a frequency of 9 192 631 770 c.p.s. A negative sign indicates that frequency is below nominal.

Arrangements have been made for the monthly publication in this Journal of the received frequencies of the above-mentioned signals as measured in this laboratory.

KALRA, S. N., BAILEY, R., and DAAMS, H. 1958. Can. J. Phys. **36**, 1442.
——— 1959. Nature (London), **183**, 375.
KALRA, S. N., PATTENSON, C. F., and THOMSON, M. M. 1959. Can. J. Phys. **37**, 10.

RECEIVED SEPTEMBER 10, 1959.
DIVISION OF APPLIED PHYSICS,
NATIONAL RESEARCH COUNCIL,
OTTAWA, CANADA.

S. N. KALRA

THE PHYSICAL SOCIETY

MEMBERSHIP of the Society is open to all who are interested in Physics.

FELLOWS pay an Entrance Fee of £1 1*s*. ($3.15) and an Annual Subscription of £2 2*s*. ($6.00).

STUDENTS: A candidate for Studentship must be between the ages of 18 and 26, and pays an Annual Subscription of 5*s*. ($0.75).

MEETINGS: Fellows and Students may attend all Meetings of the Society including the annual Exhibition of Scientific Instruments and Apparatus.

PUBLICATIONS include the *Proceedings of the Physical Society*, published monthly, and *Reports on Progress in Physics*, published annually. Volume XXI, 1958, is now available (price £3 3*s*. ($9.45)). Members are entitled to receive any of the Publications at a reduced rate.

Further information can be obtained from:

THE PHYSICAL SOCIETY
1, LOWTHER GARDENS, PRINCE CONSORT ROAD
LONDON, S.W.7, ENGLAND

NOTES TO CONTRIBUTORS

Canadian Journal of Physics

MANUSCRIPTS

General.—Manuscripts, in English or French, should be typewritten, double spaced, on paper $8\frac{1}{2}\times11$ in. **The original and one copy are to be submitted.** Tables and captions for the figures should be placed at the end of the manuscript. Every sheet of the manuscript should be numbered. Style, arrangement, spelling, and abbreviations should conform to the usage of recent numbers of this journal. Greek letters or unusual signs should be written plainly or explained by marginal notes. Characters to be set in boldface type should be indicated by a wavy line below each character. Superscripts and subscripts must be legible and carefully placed. Manuscripts and illustrations should be carefully checked before they are submitted. Authors will be charged for unnecessary deviations from the usual format and for changes made in the proof that are considered excessive or unnecessary.

Abstract.—An abstract of not more than about 200 words, indicating the scope of the work and the principal findings, is required, except in Notes.

References.—References should be listed **alphabetically by authors' names,** unnumbered, and typed after the text. The form of the citations should be that used in current issues of this journal; in references to papers in periodicals, titles should not be given and only initial page numbers are required. The names of periodicals should be abbreviated in the form given in the most recent *List of Periodicals Abstracted by Chemical Abstracts*. All citations should be checked with the original articles and each one referred to in the text by the authors' names and the year.

Tables.—Tables should be numbered in roman numerals and each table referred to in the text. Titles should always be given but should be brief; column headings should be brief and descriptive matter in the tables confined to a minimum. Vertical rules should not be used. Numerous small tables should be avoided.

ILLUSTRATIONS

General.—All figures (including each figure of the plates) should be numbered consecutively from 1 up, in arabic numerals, and each figure referred to in the text. The author's name, title of the paper, and figure number should be written in the lower left corner of the sheets on which the illustrations appear. Captions should not be written on the illustrations.

Line drawings.—Drawings should be carefully made with India ink on white drawing paper, blue tracing linen, or co-ordinate paper ruled in blue only; any co-ordinate lines that are to appear in the reproduction should be ruled in black ink. Paper ruled in green, yellow, or red should not be used. All lines must be of sufficient thickness to reproduce well. Decimal points, periods, and stippled dots must be solid black circles large enough to be reduced if necessary. Letters and numerals should be neatly made, preferably with a stencil (**do NOT use typewriting**) and be of such size that the smallest lettering will be not less than 1 mm high when the figure is reduced to a suitable size. Many drawings are made too large; originals should not be more than 2 or 3 times the size of the desired reproduction. Whenever possible two or more drawings should be grouped to reduce the number of cuts required. In such groups of drawings, or in large drawings, full use of the space available should be made; the ratio of height to width should conform to that of a journal page ($4\frac{3}{4}\times7\frac{1}{2}$ in.), but allowance must be made for the captions. **The original drawings and one set of clear copies (e.g. small photographs) are to be submitted.**

Photographs.—Prints should be made on glossy paper, with strong contrasts. They should be trimmed so that essential features only are shown and mounted carefully, with rubber cement, on white cardboard, with no space between those arranged in groups. In mounting, full use of the space available should be made. **Photographs are to be submitted in duplicate;** if they are to be reproduced in groups one set should be mounted, the duplicate set unmounted.

REPRINTS

A total of 50 reprints of each paper, without covers, are supplied free. Additional reprints, with or without covers, may be purchased at the time of publication.

Charges for reprints are based on the number of printed pages, which may be calculated approximately by multiplying by 0.6 the number of manuscript pages (double-spaced typewritten sheets, $8\frac{1}{2}\times11$ in.) and including the space occupied by illustrations. Prices and instructions for ordering reprints are sent out with the galley proof.

Contents

J. Katzman—The sun as a source of cosmic rays of intermediate energies - 1207

A. E. Covington and Gladys A. Harvey—Resolving power of three antenna patterns derived from the same aperture - - - - - - - - - - 1216

R. J. Berry—The temperature–time dependence of the triple point of water - 1230

Z. J. Kiss and H. L. Welsh—The pressure-induced rotational absorption spectrum of hydrogen - - - - - - - - - - - - - - - - 1249

P. A. Redhead—The magnetron gauge: A cold-cathode vacuum gauge - - 1260

D. R. Hay and G. E. Poaps—Frontal perturbation of a tropospheric scatter path 1272

G. R. Hanes—Limiting precision in optical interferometry - - - - - - 1283

G. Herzberg and H. R. Moore—The spectrum of Li^+ - - - - - - - 1293

W. Paskievici—Principe d'une analyse des courbes expérimentales composées: Application à la mesure de la densité des grains dans les traces fortement ionisées dans les émulsions nucléaires - - - - - - - - - - - 1314

Notes:

J. M. Daniels and M. A. R. LeBlanc—Some observations on nuclear alignment in cobalt metal - - - - - - - - - - - - - - - 1321

Irwin Oppenheim and Myer Bloom—Nuclear spin relaxation in low-temperature liquids - - - - - - - - - - - - - - - - - - 1324

Letters to the Editor:

S. N. Kalra—Frequency measurement of standard frequency transmissions against caesium-beam resonator standard - - - - - - - - 1328

Printed in Canada *by* University of Toronto Press